TÜRK MATEMATİK DERNEĞİ YAYINLARI

Sayı : 28

A. M. YAGLOM - I. M. YAGLOM

İHTİMALİYET ve INFORMASYON

Çeviren:

Lutfi BİRAN

İSTANBUL — 1966

TÜRK MATEMATİK DERNEĞİ YAYINLARI

Sayı : 28

A. M. YAGLOM - I. M. YAGLOM

İHTİMALİYET ve INFORMASYON

Çeviren:

Lutfi BİRAN

İSTANBUL — 1966

İSMAİL AKGÜN MATBAASI
Cağaloğlu, Servili Mescit Sokak, No. 24
İSTANBUL — 1966

YAZANLARIN ÖNSÖZÜ

Uzun yıllar süresince teorik matematik alanında yapılmış araştırmaların neticelerinden hiçbiri, mütehassıs matematikçilerin dar çemberinin dışına çıkmamıştır; bu keyfiyet, bazı mütehassıs olmayanları matematik'in, bu alanda araştırma yapmayı, imkânsız olmasa bile ziyadesiyle güç kılan, bir «içine çevirik» karaktere malik bulunduğunu düşünmeye sevk etti; halbuki bu böyle değildir. Bunun sebebi, matematik dergilerinde yayınlanan çalışmaların büyük bir kısmının, ziyadesiyle inkişaf etmiş bir merhaleye ulaşan matematik dalları ile ilgili bulunması ve hususî hazırlığı olmayanlar tarafından anlaşılmasının çok güç olmasındadır; elementer geometri gibi, matematik'in daha elementer alanlarına gelince, bunların, birçok yüzyıllar süren inkişafları devamınca üzerinde durulmayan problemler yahut teoremlerini bulmayı ümit etmek hayli güçtür.

Amerikalı matematikçi ve mühendis Claude SHANNON 1947 - 1948 de, inkişafı çok elementer esaslara bağlı olan, matematik'in yeni bir bölümünü keşfetti. Daha sonra «informasyon teorisi» adı verilen bu yeni bölümün menşeini, tamamen teknik mahiyette olan hat yolu yahut radyo ile ulaştırma problemleri teşkil eder. Genel bir tarzda, matemat'k'in tekniğe ve fizik bilimlerine tatbiki, modern bilimin ve tekniğin ortaya koyduğu güç problemleri incelemeden izahı mümkün olmayan karışık matematik meselelerine sıkı sıkıya bağlı bulunur; bu sebepten matematik'e ait modern araştırmaların genelleştirme imkânı pek azdır. Gene bu sebepten mütehassıs olmayanlar, matematik'in tatbikatının, kendilerine mektepte öğretildiği gibi, sadece Mısır'da arazi ölçü işlerine

inhisar ettiğini düşünürler. Bu açıdan bakılırsa, informasyon teorisi çok ilgi çekicidir, zira bu teorinin teknik problemlere tatbiki, çok az matematik kültürüne sahip okuyucular tarafından da anlaşılabilir.

Anlaşılması ancak lise seviyesinde matematik bilgisine luzum gösteren bu kitap, bir bilim disiplini olarak kabul edilen informasyon teorisine kesin bir giriş iddiasını kattiyen taşımamaktadır. Biz burada teorimin ancak, birkaç pratik tatbikatının sathî izahını yaptık; teoriye bağlı, temel teorik problemler üzerinde durulmadı. Yazanların gayesi çok basittir: okuyucunun, zamanımızın matematik'in'n basit fakat çok önemli bir kaç problemine ilgisini çekmek ve misâller üzerinde matematik metotlarının modern tekniğe uygulanması imkânlarını göstermekten ibarettir.

Kitabın birinci bölümü, sonraki bölümlerin anlaşılması için lüzumlu olan, *ihtimaliyet* kavramına tahsis edilmiştir. İkinci bölümde, SHANNON tarafından bulunmuş ve teorik ehemmiyeti son yıllarda kabul edilmiş olan *entropi* ve *informasyon* kavramları incelenmiştir; üçüncü ve dördüncü bölümler misâllere ve tatbikata tahsis edilmiştir; bu son iki bölümde, ilk iki bölümden farklı olarak, meselelerin kesin ispatları taslak halinde verilmiş ve meselelerinin kendileri de, çoğu zaman sadace çok muhtemel hipotezler gibi ifade edilmiştir. Entropi ve informasyon kavramlarının ehemmiyeti, kalp madeni paranın bulunması vs. gibi misaller yardımı ile gösterilmiştir. ve bu misaller, bazı bakımlardan, oyun kartları ve zarlarla ilgili «oyun - problemler»ini hatırlatır ve XVII inci yüzyılda ihtimaliyet teorisinin menşeini bu problemlerde aramak lâzımdır; daha teknik tatbikat misalleri dördüncü bölümde incelenmiştir.

Bu kitap matematik ile ilgilenenler ve bilhassa matematik okutanlar ve okuyanlar için hazırlanmıştır. Bunun yanında çalışmaları dolayısiyle haberleşme problemleriyle karşılaşan okuyucuları da bu kitabın ilgilendirebileceğini ümit ederiz.

BU YAYINLAR HAKKINDA

Matematiğin kendi değeri yanında, fizik, kimya ve dolayısiyle mühendislik, askerlik gibi pratik sahalara ve bilhassa son zamanlarda biyoloji, ekonomi ve hattâ sosyal bilimlere yardımı hızla arttığından, bu bilim her millet için hayatî bir önem kazanmıştır.

Öteyandan, matematik de bu bilimlerin problemlerini çözebilmek için gerek metot, gerekse fikir bakımından gelişmek zorundadır.

Bundan dolayı bir çok memleketlerde bu sahaya daha fazla sayıda istidatları çekmek ve bunları erkenden keşfetmek, en nihayet bunların eğitimi için her türlü fedakârlığa katlanmak en önemli bir millî eğitim siyaseti olmuştur.

Yeni istidatları erkenden keşfetmek için düşünülen tedbirlerin başında, matematik kültürünü geniş kitlelere yaymak gelir.

İkinci Dünya Harbinden sonra birçok memleketlerde, gençlerin tecessüslerini tahrik etmek ve bunların matematik bilimlerine karşı ilgilerini artırmak için yeni bir tip matematik literatürü meydana gelmiştir. Bu çeşit literatürde aranılan özellikler kısaca şunlar olmalıdır : a) Problem vaz'ı sun'î olmamalı, b) Bunları anlatmak için fazla ön bilgiye ihtiyaç bulunmamalı, c) Okuyucuyu aktif işbirliğine ve bir şeyler keşfetmeğe sevketmeli.

İşte Türk Matematik Derneği bu cereyanı memleketimize de getirmek maksadiyle bu yayınlara başlamış bulunmaktadır. Bu yayınlar, resmî müfredata bağlı ders veya yardımcı kitaplar olmayıp, konuları yukarıki prensiplere uygun olarak seçilmiş eserlerdir. Bunların anlaşılması için lise matematiğinin bir kısmı ile okuyucunun sağduyusu ve iyi niyeti kâfidir.

Tamamen hizmet olan bu teşebbüsümüzün malî kaynağı, Millî Eğitim Bakanlığımızın ve Ford Foundation'nun bağışlarıdır.

Türk Matematik Derneği

İÇİNDEKİLER

BÖLÜM I.
İHTİMALLER

		Sahife No.
	Yazanların Önsözü	III
	Bu yayınlar hakkında	V
1.	İhtimalin tarifi	1
2.	İhtimallerin özellikleri. Olayların toplamı ve çarpımı. Müstakil olan ve uygun olmayan olaylar	9
3.	Bağlı ihtimaller	21
4.	Olaylar cebiri ve ihtimalin genel tarifi	28

BÖLÜM II.
ENTROPİ VE İNFORMASYON

1.	Belirsizlik derecesinin bir ölçüsü olarak gözönüne alınan entropi	41
2.	Bağlı entropi	55
3.	İnformasyon anlamı	73
4.	Özellikleri ile tarif edilen entropi	82

BÖLÜM III.
BAZI LOJİK PROBLEMLERİNİN İNFORMASYON ANLAMI YARDIMI İLE ÇÖZÜMÜ

1.	Basit misâller	96
2.	Kalp madeni paraları tartı yolu ile belirleme problemleri	107

BÖLÜM IV.
HAT YOLU İLE ULAŞTIRMA PROBLEMLERİNE İNFORMASYON TEORİSİNİN TATBİKİ

1.	Temel anlamlar. Bir şifrenin ekonomisi	136
2.	Shannon - Fano şifresi. Şifrelemenin temel teoremi	147
3.	Ulaştırmada bir bulanıklığın mümkün olması hali	167

I. İLÂVE
KONVEKS FONKSİYONLARIN ÖZELLİKLERİ 183

II. İLÂVE
BAZI EŞİTSİZLİKLER 205

BÖLÜM I

İHTİMALLER

§ 1. İhtimalin tarifi

Tatbikatta, bilmediğimiz yahut hesaba katmadığımız sebeplerle değişen neticeler veren deneylere çoğu zaman rastlanır. Meselâ bir zar (altı yüzü 1 den 6 ya kadar numaralanmış homogen küp) atıldığı vakit üst yüzde hangi numaranın görüleceğini önceden bilmek imkânsızdır, zira bu, bizce meçhul olan birçok sebeplere (zarı atan elin hareketi, atıldığı esnada zarın eldeki durumu, zarın atıldığı sathın hususiyetleri vs.) tâbidir. Aynı şekilde, şu veya bu fakülteye yazılan lise mezunlarının sayısını, bir fabrikada imâl edilen parçalardan kusurlu olanların sayısını yahut, gelecek yılın yağmurlu günlerinin sayısını önceden bilmek mümkün değildir; keza bir öğrencinin bir imtihanda yapacağı hata sayısını, veyahut bir piyangoda hangi numaranın kazanacağını evvelden kestirmek imkânsızdır. Bu nevi misâllerin sayısını istediğimiz kadar artırabiliriz.

Bu tip hallere matematik'in tatbiki şu müşahedeye dayanır: birçok hallerde, aynı deney, aynı şartlar altında, birçok defa tekrarlanırsa, *gözönüne alınan neticenin sıklığı* (yani bu neticenin görüldüğü deneylerin sayısının, yapılan tekmil deneylerin sayısına oranı) *his olunacak surette daima sabit kalır* ve sabit bir p sayısına yakın bulunur. Meselâ, belirli şartlar altında atış yapan, muayyen bir atıcının yaptığı muayyen atışlarda hedefe isabet edenlerin

sayısı hemen hemen aynıdır ve bir ortalama yüzdeden çok az fark eder (bu ortalama yüzde, tabiatiyle zaman ile değişebilir: çoğalırsa, atıcının nişancılığı artıyor, eksilirse, nişancılığı azalıyor denir). Keza zar oyununda 6 nın görülmesinin sıklığı, yahut bir fabrikada belirli şartlar altında imâl edilen parçalardan kusurlu olanların yüzdesi, bunlara tekabül eden «deneyler» (zarın atılması yahut parçanın imâli) büyük sayıda tekrarlanırsa umumiyetle çok az değişir. Bu müşahedeye dayanılarak, hedefe vuran atışların, zar oyununda 6 nın görülmesinin, kusurlu bir parçanın imâlinin, vs., her hal için belirli bir sayı ile karakterize edilebileceği neticesine varılır. İşte bu sayıya gözönüne alınan olayın *ihtimali* denir; büyük sayıda deneyler içinde uygun neticelerin sıklığı bu ihtimale yakın olacaktır. Matematik, mekanik, fizik, bioloji ile ilgili diğer birçok meselelerde, ihtimal, tamamen benzer tarzda tarif edilir. İhtimalleri ve bunların tatbikatını inceleyen bilime *ihtimaller teorisi* denir.

Yukarıda verdiğimiz kısa izaha göre, bir olayın ihtimali, uzun bir deney serisinin neticesinden, hesaplanabilir. Bununla beraber, bir ihtimalin mevcudiyetinin bu deneylerin yapılmasına hiçbir tarzda bağlı olmadığı aşikârdır. Bu takdirde, muayyen olayların ihtimallerinin, bu deney serisine başvurmadan, hangi metodlarla belirlenebileceğinin mümkün olacağı sorulabilir; bir defa ihtimalin bu belirelnmesi yapıldıktan sonra, yapılması düşünülen deneylerin neticesini evvelden kestirmek mümkün olacak ve bu, ihtimal anlamının geniş bir sahada fiziksel tatbikatına imkân verecektir; bu meseleyi burada bütün teferruatiyle incelemeyeceğiz; sadece çok basit bir misâl ile iktifa edeceğiz, fakat birçok benzer problemi bu misâl yardımı ile çözmek mümkündür.

İçerisinde, ancak renk bakımından birbirinden farklı

on tane top bulunan bir sandık gözönüne alalım. Toplardan 5 i beyaz, 3 ü siyah ve 2 si kırmızı olsun ve tamamen karıştırılmış bulunsun. Bakmadan sandıktan bir top çıkaralım; bu topun bu üç renkten birinde olması ihtimali soruluyor. Şurası aşikâr ki, bir beyaz top çıkarılması şansı 10 üzerine 5, bir siyah top çıkarılması şansı 10 üzerine 3, ve bir kırmızı top çıkarılması ihtimali 10 üzerine 2 dir; başka ifade ile, bir beyaz top, bir siyah top ve bir kırmızı top çıkarılmasının ihtimalleri, sıra ile, $5/10 = 1/2$, $3/10 =$ ve $2/10 = 1/5$ dir. Gerçekte, deney birçok (çıkarılan topu her defasında sandığa koyarak ve iyice karıştırarak) tekrarlanırsa, deneylerin yaklaşık olarak % 50 sinde beyaz top, % 30 unda siyah top ve % 20 sinde kırmızı top çıkacaktır. Tabiatiyle problem, herhangi sayıda farklı renklerle, aynı tarzda çözülecektir.

Şimdi, «sandık problemi» yolu ile ihtimallerin belirlenmesine ait birkaç misâl verelim.

PROBLEM 1. — Gelişigüzel atılmış bir madenî parada «tura» gelmesi ihtimali nedir?

Bu problem, aşikâr olarak, aşağıdakinin aynıdır: bir sandık içerisinde, birisinin üzerinde «tura», diğerinin üzerine «yazı» işareti bulunan iki top bulunsun (tabiatiyle bu işaretler yerine iki farklı renk meselâ beyaz ve siyah alınabilir), sandıktan gelişigüzel çıkardığımız bir topun «tura» işaretli olması ihtimali nedir? Aranılan ihtimalin 1/2 olduğu aşikârdır.

PROBLEM 2. — Bir zar atıldığı vakit 3 ile bölünebilen bir sayı gelmesi ihtimali nedir?

Zarın atılması yerine, 1, 2, 3, 4, 5, 6 numaralarını taşıyan altı topun bulunduğu bir sandıktan bir top çıkarıl-

masını alabiliriz. 3 ve 6 numaralı topları siyaha ve diğerlerini beyaza boyarsak, mesele, sandıktan çıkarılan bir topun siyah olması ihtimalinin belirlenmesine getirilir (gerçekten, 3 ve 6 rakamları 3 ile bölünebilir diğerleri bölünemez). Bu şartlar altında aranılan ihtimalin 2/6=1/3 olduğu aşikârdır.

PROBLER 3. — Bir öğrenciler toplantısına bir okuldan 200, ikinci bir okuldan 250 ve bir üçüncüsünden 300 öğrencinin katıdığı biliniyor. Tesadüfi olarak görüştüğümüz bir öğrencinin ikinci okul öğrencisi olması ihtimali nedir?

Bu problem şuna denktir: içerisinde 750 top bulunan bir sandık gözönüne alalım; bunların 200 ü beyaz, 250 si siyah ve 300 ü kırmızı olsun. Gelişigüzel çekilen bir topun siyah olması ihtimali nedir? Bu ihtimalin 250/750 = 1/3 olduğu aşikârdır.

Şimdi bu problemlerin çözülmesini mümkün kılan genel prensipleri anlamaya çalışalım. Problemlerden evvel incelenen misâlde, topların iyice karıştırıldığını ve top çekilirken bakılmadığını farzettik; bu, sandıktan çıkarılma şansının her top için aynı olması veyahut sandıktan çıkarılmanın her top için *eşit olarak muhtemel* olması demektir. Sandıkta 10 top bulunduğundan her top için ihtimal 1/10 dur. Diğer taraftan, 5 beyaz top olduğundan bir beyaz topun çekilmesi ihtimali 5/10 dur.

Benzer düşünceler, diğer üç problemin her birinin çözülmesini mümkün kılar. Meselâ, bir zar atıldığı vakit zarın 6 yüzünden herhangi birisinin gelmesinin eşit olarak muhtemel olduğunu gözönüne aldık, ve bu sebepten dolayıdır ki, bu problemin yerine, içerisinde altı top bulunan sandık problemini alabildik. Fakat 6 yüzden ancak ikisinde

aranılan şartları gerçekleyen sayı bulunmaktadır; zarın bu 2 yüzünden birinin gelmesi ihtimali 2/6 dır.

İncelenen deneyin (bir topun çekilmesi, bir madenî paranın veya bir zarın atılması, ikinci okuldan bir öğrenci ile görüşme vs.), eşit olarak muhtemel olan n sayıda neticesi var ise, bu neticelerden her birinin ihtimali $1/n$ ye eşit olacaktır. Şimdi deneyin neticesi ile belirli bir olay (bir beyaz topun çekilmesi, düşen paranın «tura» gelmesi atılan zarda çift sayı gelmesi, görüşülen öğrencinin ikinci okuldan olması, vs.) gözönüne alalım. Bir olay, m hal içinde n defa vuku bulur ve $m - n$ defa vuku bulmazsa, bu olayın ihtimali n/m ye eşittir denir. Başka ifade ile, bir olayın ihtimali, eşit olarak muhtemel olan uygun neticelerin sayısının, eşit olarak muhtemel olan tekmil neticelerin sayısına olan oranına eşittir. Bu son ifade, bir olayın ihtimalinin tarifi olarak alınabilir. Mümkün neticelerin eşit olarak muhtemel olması keyfiyeti, yapılan deneyde açık olarak belirtilmelidir (meselâ zarın tamamen küp şeklinde ve homogen maddeden yapılmış olması, topların tamamen karıştırılmış ve birbirlerinden ancak renk bakımından farklı bulunması gibi). İhtimallerin belirlenmesine ait bazı mühim haller bu tarifin dışında kalmasına rağmen, bizim için bu tarif tamamen kâfi gelecektir.

Şimdi bu kitapta kullanacağımız terimleri tespit edelim. Bir deney neticesinde vuku bulabilen veya vuku bulamayan bir olaya *tesadüfî olay* diyeceğiz. Bir deneyin *neticesine* aynı mânayı vereceğiz. Tesadüfi olaylar majüskül italik harflerle, bunların ihtimali (yahut bazı belirli neticelerin ihtimali) p harfi ile gösterilecektir; A olayının ihtimali çoğu zaman $p(A)$ ile gösterilecektir. Birçok farklı neticeleri olabilen deneyler bizim için mühim rol oynıyacaktır; tekmil bu neticeleri, her birine farklı indisler koyarak aynı harfle göstereceğiz; deneylerin kendileri,

yunan harfleri ile gösterilecektir. Bu tipte her deneye bir ihtimal cedveli tekabül ettirilebilir:

deney neticeleri	A_1	A_2	...	A_k
ihtimaller	$p(A_1)$	$p(A_2)$	...	$p(A_k)$

Meselâ, birinci misâlimize şu cetvel tekabül eder:

A_1	A_2	A_3
$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{5}$

(burada A_1, bir beyaz top çekilmesini A_2, bir siyah top çekilmesini ve A_3, bir kırmızı top çekilmesini gösterir.

Problem 1 e şu cetvel tekabül eder:

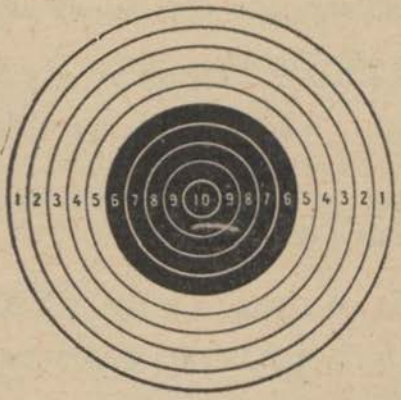
B_1	B_2
$\frac{1}{2}$	$\frac{1}{2}$

(burada B_1 «tura», B_2 «yazı» gösterir). Zarın atılmasına şu cetvel tekabül eder:

Yüzün numarası	1	2	3	4	5	6
ihtimal	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Bu son cetvel ile ilk iki cetvel arasında esaslı bir farka işaret etmek lâzımdır. Sonucu da, deney neticesi (1, 2, 3,

4, 5, 6 gibi) bir sayı ile gösterilebilir, bu, diğer iki halde mümkün değildir. Zarın atılması ile elde edilen sayı (bizim kontrolumuzun dışındaki sebeplere göre), altı mümkün değer alabilen bir *tesadüfî büyüklük* teşkil eder denebilir. Tesadüfî büyüklükler için başka misâller verilebilir: imâl edilmiş yüz parçadan kusurlu parçaların sayısı, muayyen bir şehirde bir senede doğum sayısı, bir atıcı tarafından belirli şartlar altında bir hedefe yapılan atışlarda elde edilen sayı miktarı (Şekil 1 de, her bölge için elde edilecek sayı, işaretlenmiş bir hedef gösetrir) vs. *.



Şekil: 1

«Tesadüfî büyüklük» anlamı, bu büyüklüğün değerini, şu veya bu tarzda değerlendirilmesini icap ettirir. Bunun nasıl yapılacağını anlamak kolaydır. Meselâ, hemen yukarıda verilen misâllerden birincisini (imâl edilmiş yüz parçadan kusurlu olanların sayısını) gözönüne alalım; belirli imâl şartarı altında bu sayının 6 yı aşmadığını farz edelim, ve ihtimaller cetveli aşağıdaki şekilde olsun:

---

*) Bu kitabın konusu dışında kalan tesadüfî büyüklük anlamı, ihtimaller teoresinde çok mühim bir rol oynar.

Kusurlu parçaların sayısı	0	1	2	3	4	5	6
İhtimaller	0,1	0,15	0,2	0,25	0,15	0,1	0,05

Bu şartlar altında, N ile göstereceğimiz büyük bir sayıda imâl edilmiş muayyen parçalardan yüzde $0{,}1\,N$ kadarında kusurlu parça olmayacak, yüzde $0{,}15\,N$ kadarında bir kusurlu parça bulunacak, yüzde $0{,}2\,N$ kadarında iki kusurlu parça bulunacak, $0{,}25\,N$ de üç, $0{,}15\,N$ de dört, $0{,}1\,N$ de beş ve $0{,}05\,N$ de altı kusurlu parça olacaktır. Bunun neticesi olarak, N büyük ise, kusurlu parçaların toplamı olarak,

$$a = 0{,}1\,N \cdot 0 + 0{,}15\,N \cdot 1 + 0{,}2\,N \cdot 2 + 0{,}25\,N \cdot 3 + \\ + 0{,}15\,N \cdot 4 + 0{,}1\,N \cdot 5 + 0{,}05\,N \cdot 6$$

alınabilir ve imâl edilmiş yüz parçadan kusurlu parçaların sayısının ortalama değeri

$$\frac{a}{N} = 0{,}1 \cdot 0 + 0{,}15 \cdot 1 + 0{,}2 \cdot 2 + 0{,}25 \cdot 3 + \\ + 0{,}15 \cdot 4 + 0{,}1 \cdot 5 + 0{,}05 \cdot 6 = 2{,}7$$

olacaktır. Daha genel olarak tesadüfî büyüklüğün ihtimallerinin cetveli

Tesadüfî büyüklüğün değeri	a_1	a_2	a_3	...	a_k
İhtimaller	p_1	p_2	p_3	...	p_k

dir; ortalama değeri ise şu olacaktır:

$$a \text{ nın ort. değeri} = p_1 \cdot a_1 + p_2 \cdot a_2 + p_3 \cdot a_3 + \cdots + p_k \cdot a_k .$$

PROBLEM 4. — A ve B gibi iki atıcının hedefe atış için ihtimal cetvelleri şunlardır:

Atıcı A :

sayı miktarı	0	1	2	3	4	5	6	7	8	9	10
ihtimaller	0,2	0,03	0,05	0,1	0,15	0,2	0,1	0,1	0,07	0,05	0,03

Atıcı B :

sayı miktarı	0	1	2	3	4	5	6	7	8	9	10
ihtimaller	0,01	0,01	0,04	0,1	0,25	0,3	0,18	0,05	0,03	0,02	0,01

Bu iki atıcıdan hangisi daha nişancıdır?

Bu halde A atıcısı arafından elde edilen sayıların ortalama değeri

$$0{,}02\cdot 0+0{,}03\cdot 1+0{,}05\cdot 2+0{,}1\cdot 3+0{,}15\cdot 4+0{,}2\cdot 5+$$
$$+0{,}2\cdot 6+0{,}1\cdot 7+0{,}07\cdot 8+0{,}05\cdot 9+0{,}03\cdot 10=5{,}24$$

B atıcısı için ise

$$0{,}01\cdot 0+0{,}01\cdot 1+0{,}04\cdot 2+0{,}1\cdot 3+0{,}25\cdot 4+0{,}3\cdot 5+0{,}18\cdot 6+$$
$$+0{,}05\cdot 7+0{,}03\cdot 8+0{,}02\cdot 9+0{,}01\cdot 10=4{,}84<5{,}25.$$

Şu halde birincisi daha nişancıdır.

§ 2. İhtimallerin özellikleri. Olayların toplamı ve çarpımı. Müstakil olan ve uygun olmayan olaylar

Evvelki paragrafta verilen ihtimal tarifi mucibince, herhangi bir A olayının $p(A)$ ihtimali, iki tam sayının bölümü olup

$$0 \leq p(A) \leq 1$$

dir. Bu ihtimal 1 e eşit olabilir: bu deney neticesi *ne olursa olsun* A olayı vuku bulacaktır demektir; başka ifade ile, A olayı *kesin*dir (meselâ, içinde yalnız beyaz toplar bulunan bir sandıktan bir beyaz top çekme ihtimali 1 dir). İhtimal sıfıra eşit olabilir: bu ise, deney neticesi *ne olursa olsun* olay asla vuku bulmaz demektir; başka ifade ile, A olayı *imkânsız*dır (içinde yalnız beyaz toplar bulunan bir sandıktan bir siyah top çekme ihtimali sıfıra eşittir).

Şimdi, karşılıklı olarak birleşemeyen A ve B gibi ancak iki neticesi olan bir deney gözönüne alalım. Bu halde B ye, A olayının *zıt*'ı denir ve $\overline{A}$ yazılır. Eğer A olayı, eşit olarak vukuu mümkün olan n halde m defa vuku bulursa, $\overline{A}$ olayı diğer $n-m$ halde vuku bulacaktır; bunun neticesi olarak, $p(A)=m/n$, $p(\overline{A})=(n-m)/n=1-m/n$ ve

$$p(\overline{A})=1-p(A)$$

dır. Böyle bir deney için ihtimaller cetveli şöyledir:

A	$\overline{A}$
$p(A)$	$1-p(A)$

Şimdi A ve A_1 gibi iki olay gözönüne alalım, şöyle ki A olayı A_1 olayının vukuunu icap ettirsin (meselâ, zar oyununda A, 6 gelmesi, A_1 ise 3 ile bölünebilen bir sayı gelmesidir). A nın vuku bulduğu tekmil hallerde A_1 olayı vuku bulacağından A_1 olayının ihtimali A nınkinden küçük olamaz. Bu keyfiyet $A \subset A_1$ yazılır ve şu mühim özellik elde edilir:

eğer $A \subset A_1$ ise $p(A) \leq p(A_1)$ olur.

Şimdi, ancak A ve B olaylarından biri vuku bulduğu takdirde vuku bulan bir olay gözönüne alalım; böyle bir olaya A ve B olaylarının *toplamı* denir ve $A + B$ ile gösterilir. Bu takdirde iki hal mevcut olabilir. Eğer A ve B olayları *uyuşamaz* ise, yani karşılıklı olarak birleşmezlerse, A, eşit olarak mümkün olan n olayın m_1 inde ve B, diğer m_2 halde vuku bulur ve

$$p(A) = \frac{m_1}{n}, \quad p(B) = \frac{m_2}{n} \quad \text{ve}$$

$$p(A + B) = \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n}$$

buradan,

$$p(A + B) = p(A) + p(B)$$

(toplam ihtimaller kanunu) bulunur. Böylece sahife 2-3 de incelenen misâlde, bir beyaz veya siyah top çekilmesi ihtimali, toplam ihtimaller kanunu mucibince

$$\frac{1}{2} + \frac{3}{10} = \frac{4}{5}$$

e eşittir.

Toplam ihtimaller kanunu ile ifade olunan ihtimallerin özellikeri aşağıdaki tarzda genelleştirilebilir. İkişer ikişer birleşemeyen k tane $A_1, A_2, \cdots, A_k$ olaylarını gözönüne alalım. Bu k olaydan hiç olmazsa biri vuku bulduğu vakit vuku bulan olayı $A_1 + A_2 + \cdots + A_k$ ile gösterelim. Bu takdirde

$$p(A_1 + A_2 + \cdots + A_k) = \\ = p(A_1) + p(A_2) + \cdots + p(A_k)$$

olacağı aşikârdır. Daha genel olan bu neticeye de toplam ihtimaller kanunu denir. Hususi olarak, k (ve yalnız k) sayıda ve karşılıklı olarak birleşemeyen neticesi olan bir deneyde,

A_1	A_2	$\cdots$	A_k
$p(A_1)$	$p(A_2)$	$\cdots$	$p(A_k)$

ihtimaller cetvelinde ikinci satırın sayılarının toplamı 1 e eşittir:

$$p(A_1) + p(A_2) + \cdots + p(A_k) = 1;$$

bu,

$$p(A_1) + p(A_2) + \cdots + p(A_k) = \\ = p(A_1 + A_2 + \cdots + A_k)$$

eşitliğinin ve $A_1 + A_2 + \cdots + A_k$ olayının kesin olmasının neticesidir (deney daima mümkün olan neticelerden hiç olmazsa birini verir).

Şimdi, A ve B olaylarının *birleşebildiklerini,* yani aynı zamanda vuku bulabileceklerini farzedelim. Artık $p(A + B) = p(A) + p(B)$ olduğu iddia edilemez. Zira, n tane eşit olarak mümkün olan hallerin m_1 tanesinde A vuku bulur ve m_2 tanesinde B vuku bulursa, $A + B$ olayı, m_1 neticelerinin veya m_2 neticelerinin birinde vuku bulur; bununla beraber bu neticeler *artık mutlak suretle farklı*

olmadıklarından, farklı neticeelrin toplam sayısı $m_1 + m_2$ den küçük olabilir. Demek ki bu halde sadece

$$p(A + B) \leq p(A) + p(B)$$

olduğu söylenebilir (bununla beraber $p(A + B) \geq p(A)$ ve $p(A + B) \geq p(B)$ dir, zira olayların toplamının tarifi mucibince $A \subset A + B$ ve $B \subset A + B$ dir).

Bu son eşitsizliği açıklayalım: A ve B gibi iki olay vuku bulduğu takdirde vuku bulan olaya bu iki olayın çarpımı adını verelim ve bunu AB ile gösterelim. A ya uygun ve eşit olarak mümkün olan m_1 neticeleri, ve B ye uygun ve eşit olarak mümkün olan m_2 neticeleri arasında l tane identik neticenin bulunduğunu farz edelim. Bu l neticeden biri vuku bulursa (ve yalnız bu halde), A ve B olayları vuku bulur; bunlardan $p(AB) = l/n$ neticesi çıkarılır. Bundan başka, ilk m_1 netice ve ikinci m_2 netice arasında identik neticeler bulunduğundan hepsi $m_1 + m_2 - l$ sayıda farklı netice vardır ($m_1 + m_2$ toplamında, l netice iki defa sayılmıştır). Şu halde

$$p(A + B) = \frac{m_1 + m_2 - l}{n} = \frac{m_1}{n} + \frac{m_2}{n} - \frac{l}{n}$$

ve

$$p(A + B) = p(A) + p(B) - p(AB)$$

olur. Şu halde $A + B$ olayının ihtimalinin A ve B olaylarının ihtimalleri cinsinden belirlenmesi, A ve B olaylarının AB çarpımının ihtimalinin belirlenmesine getirilir. Genel halde basit olmayan bu mesele gelecek paragrafta incelenecektir. Bununla beraber, AB olayının ihtimalinin belir-

lenmesi güçlük göstermeyen bir hal vardır. Bu, A ve B olaylarının müstakil olması halidir; başka ifade ile, bu, A olayının vuku bulmasına bağlı neticenin, B nin vuku bulmasının bağlı bulunduğu deneyin şartlarına hiçbir şekilde tesir etmemesi halidir; meselâ, içlerinde beyaz ve siyah toplar bulunan *iki farklı sandıktan* siyah toplar çekmek müstakil olaylar teşkil eder; buna karşılık, *aynı sandıktan* (çekilen topu sandığa koymadan) üstüste iki top çekme müstakil olaylar olamaz (zira, ilk çekiliş sandıkta kalan top sayısına tesir eder ve deneyin şartlarını değiştirir).

A olayının, ilk deneyin eşit olarak mümkün olan n_1 neticesinin m_1 tanesinde müstakil olarak vuku bulduğunu, B olayının ise ikinci deneyin eşit olarak mümkün olan n_2 neticesinin m_2 tanesinde vuku bulduğunu farzedelim; A nın ihtimali m_1/n_1, B ninki m_2/n_2 dir. Şimdi ilk ikisinin üstüste gelmesinden husule gelen bir karma deney gözönüne alalım. Bu karma deneyin, birincinin n_1 deneyinin, ikincinin n_2 deneyinin her biri ile birleşmesiyle meydana gelen eşit olarak mümkün n_1n_2 tane neticesi olacağı aşikârdır. Bu n_1n_2 mümkün deney içinde m_1m_2 sayıda AB ye uygun deney olacaktır, bunlar ilk deneyde A ya uygun olan m_1 sayıda neticeler, ikinci deneyde B ye uygun olan m_2 sayıda neticelerle birleştirilerek elde edilir. Şu halde AB nin ihtimali $m_1m_2/n_1n_2 = (m_1/n_1) \cdot (m_2/n_2)$ olacak ve

$$p(AB) = p(A)p(B)$$

(birleşik ihtimaller kanunu) elde edilecektir. Bu kanun aşağıdaki tarzda genelleştirilebilir. $A_1, A_2, \cdots, A_k$ gibi k sayıda olay gözönüne alalım ve bunların aralarında müstakil olduklarını, yani bu olaylardan herhangi birisinin vuku bulmasına tesir eden şartların, aralarından herhangi

birisinin vuku bulmasıyla değişmediklerini farzedelim. Bu halde

$$p(A_1A_2 \cdots A_k) = p(A_1)p(A_2) \cdots p(A_k)$$

olur. Bu bağıntının ispatı, $p(AB) = p(A)p(B)$ formülünün ispatına benzer tarzda yapılır ve esasen bu son formül yukardakinin bir özel halidir.

A ve B *müstakil değilse,* birleşik ihtimaller kanunu artık gerçekleşmez. Şu halde $B \subset A$ ise (meselâ A, zar oynunda bir çift sayı gelme olayı ve B, dört gelme olayı ise), AB olayı B olayı ile üstüste gelir ve $p(AB) = p(B)$ dir. Şimdilik sadece $p(AB) \leq p(A)$ ve $p(AB) \leq p(B)$ olduğunu söyleyebiliriz (gerçekten, olaylar çarpımının tarifi mucibince $AB \subset A$ ve $AB \subset B$ dir). Gelecek paragrafta iki olayın çarpımının ihtimalini daha etraflı olarak inceleyeceğiz.

Bu çok basit özellikleri meydana çıkarmak için birkaç problem vereceğiz.

PROBLEM 5. — Yazı-tura oyununda üstüste iki defa «tura» elde etme ihtimali nedir?

Burada, A, paranın ilk atılışında «tura» gelmesi olayı ve B, ikinci atılışında «tura» gelmesi olayı olmak üzere, AB olayının ihtimali aranıyor. A ve B olayları müstakil olduklarından

$$p(AB) = p(A)p(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

dür (bak. problem 1, sahife 3).

PROBLEM 6. — 1000 den küçük gelişigüzel seçilmiş

bir tam sayının, diğer bir tam sayının tam kuvveti (üssü birden büyük) olması ihtimali nedir?

$$2^9 < 1000 < 2^{10},\quad 3^6 < 1000 < 3^7,\quad 5^4 < 1000 < 5^5,$$
$$6^3 < 1000 < 6^4,\quad 7^3 < 1000\,7^4,\quad 10^3 = 1000 < 10^4,$$
$$11^2 < 100 < 11^3,\quad 12^2 < 1000 < 12^3,\ \cdots$$
$$\cdots,\ 31^2 < 1000 < 31^3,\quad 32^2 < 1000$$

olduğundan bu sayının 2 nin bir kuvveti olması ihtimali 8/1000 dir (1 den 1000 e kadar 1000 sayıda 2 nin 8 tane kuvveti vardır: $2^2 = 4$, $2^3 = 8$, 2^4, 2^5, 2^6, 2^7, 2^8, 2^9); keza, bu sayının 3, 5, 6, 7, 10, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 26, 28, 29, 30, 31 in bir tam kuvveti olması ihtimalleri, sıra ile 5/1000, 3/1000, 2/1000, 2/1000, 2/1000, 1/1000, 1/1000, ···, 1/1000 dir (eğer sayı, 4, 8, 9, 16, 25 yahut 27 nin bir tam kuvveti ise aynı zamanda daha küçük bir sayının da tam kuvvetidir). Mütekabil olaylar ikişer ikişer birleşmediklerinden aranılan ihtimal

$$\frac{8}{1000}+\frac{5}{1000}+\frac{3}{1000}+\frac{2}{1000}+\frac{2}{1000}+\frac{2}{1000}+$$
$$+\underbrace{\frac{1}{1000}+\frac{1}{1000}+\cdots+\frac{1}{1000}}_{18\text{ defa}}=\frac{40}{1000}=\frac{1}{25}$$

e eşittir.

PROBLEM 7. — Bir zar atıldığı vakit iki ile yahut üç ile bölünebilen bir sayı gelmesi ihtimali nedir?

2 nin katı bir sayı gelmesi olayını A ile, 3 ün katı bir sayı gelmesi olayını B ile gösterelim; şu halde 6 nın katı bir sayı gelmesi olayı, AB olacaktır; bu şartlar altında

$p(A) = 3/6 = 1/2$, $p(B) = 2/6 = 1/2$, $p(AB) = 1/6$ dır (bak. problem 2, sahife 3). Buradan, aranılan ihtimal için

$$p(A+B) = p(A) + p(B) - p(AB) =$$

$$= \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{2}{3}$$

bulunur.

PROBLEM 8. — Altı avcı bir tilki görürler ve aynı zamanda ateş ederler. Aynı uzaklıktan ateş eden avcılardan her biri üç atıştan birinde isabet ettirdiği ve tilkiyi öldürdüğü farzedildiğine göre tilkinin öldürülmesi ihtimali nedir?

$A_1, A_2, \cdots, A_6$ ile, birinci, ikinci, ..., altıncı avcının tilkiyi öldürmesi olaylarını gösterelim. Hipotezlere göre $p(A_1) = p(A_2) = \cdots = p(A_6) = 1/3$ dür; $S = A_1 + A_2 + \cdots + A_6$ olduğuna göre $p(S)$ aranıyor. $A_1, A_2, \ldots, A_6$, tabiatiyle müstakil olaylardır; bu bize, problemi çözmek için,

$$p(A+B) = p(A) + p(B) - p(AB) =$$

$$= p(A) + p(B) - p(A)p(B)$$

formülünü kullanmayı mümkün kılar (biraz ileride genel hale bak). Bununla beraber, bu çözüm pek basit değildir, zira uygun olmayan birçok olayların toplamının ihtimalini ifade eden formül kâfi derecede karışıktır.

Daha çok basit olan bir çözüm vardır. Tilkinin öldürülmemesi ihtimalini arıyalım, bu ihtimal $p(\overline{S})$ olsun. Birinci, ikinci, ..., altıncı avcı tarafından yapılan isabetsiz

atışlar, sıra ile, $\overline{A}_1, \overline{A}_2, \cdots, \overline{A}_6$ dır. $p(\overline{A}) = 1 - p(A)$ olduğundan, $p(\overline{A}_1) = p(\overline{A}_2) = \cdots = p(\overline{A}_6) = 2/3$ dir. Tilkinin öldürülmemesi için tekmil avcıların isabetsiz atış yapmaları, başka ifade ile, $\overline{A}_1, \overline{A}_2, \ldots, \overline{A}_6$ olayları müstakil olmak üzere $\overline{A}_1 \overline{A}_2 \ldots \overline{A}_6$ olayının vuku bulması lâzımdır. Buradan

$$p(\bar{S}) = p(\overline{A}_1 \overline{A}_2 \cdots \overline{A}_6) = p(\overline{A}_1) p(\overline{A}_2) \cdots p(\overline{A}_6) =$$

$$= \frac{2}{3} \cdot \frac{2}{3} \cdots \frac{2}{3} = \frac{2^6}{3^6} = \frac{64}{729},$$

ve $p(A) = 1 - p(A)$ formülü vasıtasiyle

$$p(S) = 1 - \frac{64}{729} = \frac{665}{729}$$

elde edilir.

$p(A + B) = p(A) + p(B) - p(AB)$ formülü herhangi k sayıda $A_1, A_2, \cdots, A_k$ gibi olayların toplamına teşmil edilebilir (bu olayların ikişer ikişer uygun olmaması icap etmez).

$$p(A_1 + A_2) = p(A_1) + p(A_2) - p(A_1 A_2)$$

olmak üzere

$$p(A_1 + A_2 + A_3) = p\{(A_1 + A_2) + A_3\}$$

$$= p(A_1 + A_2) + p(A_3) - p\{(A_1 + A_2) A_3\}$$

yazılır. $p\{(A_1 + A_2) A_3\}$ ün mânası kolayca açıklanır.

Gerçekten, olaylar çarpımının tarifi mucibince $(A_1+A_2)A_3$ ifadesi A_1 yahut A_2 olaylarından birinin vuku bulmasını, üstelik A_3 olayının da vuku bulmasını gösterir. Bu aynı zamanda A_1A_3 ve A_2A_3 olaylarından hiç olmazsa birinin vuku bulması demektir, şu halde

$$(A_1 + A_2)A_3 = A_1A_3 + A_2A_3$$

dür, bunun neticesi olarak

$$p\{(A_1 + A_2)A_3\} = p(A_1A_3 + A_2A_3)$$
$$= p(A_1A_3) + p(A_2A_3) - p\{(A_1A_3)(A_2A_3)\}$$

yazılır. $(A_1A_3)(A_2A_3)$ olayı A_1A_3 (yani A_1 ve A_2 nin) ve A_2A_3 (yani A_2 ve A_3 ün) aynı zamanda vuku bulmasını gösterir; başka ifade ile, $(A_1A_3)(A_2A_3)$ olayı A_1, A_2 ve A_3 ün aynı zamanda vuku bulmasını gösterir ve şu halde $A_1A_2A_3$ den farksızdır.

Aynı tarzda

$$p(A_1 + A_2 + A_3) =$$
$$= p(A_1) + p(A_2) - p(A_1A_2) + p(A_3) - p(A_1A_3) -$$
$$- p(A_2A_3) + p(A_1A_2A_3)$$

yahut, diğer bir sırada yazarak,

$$p(A_1 + A_2 + A_3) =$$
$$= p(A_1) + p(A_2) + p(A_3) - p(A_2A_3) - p(A_3A_1) -$$
$$- p(A_1A_2) + p(A_1A_2A_3)$$

bulunur.

Keza k herhangi bir sayı olmak üzere

$$p(A_1 + A_2 + \cdots + A_k) =$$
$$= p(A_1) + p(A_2) + \cdots + p(A_k) - p(A_1A_2) -$$
$$- p(A_1A_3) - \cdots - p(A_{k-1}A_k) + p(A_1A_2A_3) +$$
$$+ p(A_1A_2A_4) + \cdots + p(A_{k-2}A_{k-1}A_k) -$$
$$- p(A_1A_2A_3A_4) - \cdots + (-1)^{k-1}p(A_1A_2 \cdots A_k)$$

elde edilir. Bu formül *rekürans* yolu ile ispat edilir.

Şimdi, bu formül yardımı ile problem 8 in nasıl çözüleceğini gösterelim. $k = 6$ için

$$p(A_1 + A_2 + \cdots + A_6) =$$
$$= p(A_1) + p(A_2) + \cdots + p(A_6) - p(A_1A_2) -$$
$$- p(A_1A_3) - \cdots - p(A_5A_6) + p(A_1A_2A_3) +$$
$$+ p(A_1A_2A_4) + \cdots + p(A_4A_5A_6) - \cdots -$$
$$- p(A_1A_2A_3A_4A_5A_6)$$

dır; fakat $A_1, A_2 \cdots, A_6$ olayları müstakil olduklarından

$$p(A_1) = p(A_2) = \cdots = p(A_6) = \frac{1}{3},$$

$$p(A_1A_2) = p(A_1A_3) = \cdots = p(A_5A_6) =$$
$$= p(A_1)p(A_2) = (\tfrac{1}{3})^2,$$

$$p(A_1A_2A_3) = p(A_1A_2A_4) = \cdots = p(A_4A_5A_6) =$$
$$= p(A_1)p(A_2)p(A_3) = (\tfrac{1}{3})^3,$$

$$\cdots, p(A_1A_2 \cdots A_6) = p(A_1)p(A_2) \ldots p(A_6) = (\tfrac{1}{3})^6,$$

yazılır. Buradan

$$p(A_1 + A_2 + \ldots + A_6) = 6 \cdot \frac{1}{3} - C_6^2 \left(\frac{1}{3}\right)^2 + C_6^3 \left(\frac{1}{3}\right)^3 -$$

$$- C_6^4 \left(\frac{1}{3}\right)^4 + C_6^5 \left(\frac{1}{3}\right)^5 - C_6^6 \left(\frac{1}{3}\right)^6 = 1 - \left(1 - \frac{1}{3}\right)^6 = 1 - \left(\frac{1}{3}\right)^6 = \frac{665}{729}$$

elde edilir; bu ise evvelki neticenin aynıdır.

§ 3. Bağlı ihtimaller

A ve B gibi iki olay gözönüne alındığı vakit A ya uygun bir deney neticesi B ye bağlı deney şartları üzerine hiçbir tesir yapmadığı takdirde A ve B ye iki müstakil olay demiştik. Bu hal, genel olmaktan uzaktır. Daha evvel verdiğimiz bir misâli şimdi daha etraflı olarak inceleyeceğiz. m sayıda siyah top ve $n - m$ sayıda beyaz top ihtiva eden bir sandıktan bir siyah top çekilmesi olayı A olsun; ilk topun çekilişinden sonra aynı sandıktan yine bir siyah top çekilmesi olayını da B ile gösterelim. Eğer ilk çekilen top siyah ise, yani A olayı vuku bulursa, sandıkta $m - 1$ sayıda siyah top ve $n - m$ sayıda beyaz top kalacak ve B olayının ihtimali $(m-1)/(n-1)$ olacaktır.

Eğer ilk çekilen top beyaz ise ($\bar{A}$ olayı vuku bulursa), sandıkta m sayıda siyah top ve $n - m - 1$ sayıda beyaz top kalacaktır; bu takdirde B olayının ihtimali $m/(n-1)$ olacaktır. Şu halde B nin ihtimali, A nın vuku bulup bulmamasına bağlıdır; demek ki, B nin ihtimali iki farklı değere malik olabilir: $(m-1)/(n-1)$ ve $m/(n-1)$.

A nın vuku bulduğunun bilinmesi halinde B olayının ihtimaline, *B olayının, A nın vukuuna bağlı ihtimali* denecek ve $p_A(B)$ ile gösterilecektir. Bahis konusu olan

misâlde bu ihtimal $p_A(B) = (m-1)(n-1)$ dir. Keza B nin, $\overline{A}$ nın vukuuna bağlı ihtimali (yani B nin, A vuku bulmaması halindeki ihtimali) tarif edilir ve bu $p_A(B)$ ile gösterilir; burada da $p_{\bar{A}}(B) = m(n-1)$ dir.

Bir A olayının ihtimali, bu olayın $p(B)$ ihtimalinden (yani A nın vuku bulup bulmadığı hususunda hiç bir şey bilinmemesi halinde, B nin ihtimalinden) büyük veya küçük olabilir. Şu halde, misâlimizde, sandıkta bulunan n toptan herhangi birisinin çekilmesi şansı aynı olduğundan ve m tane siyah top mevcut bulunduğundan $p(B) = m/n$ dir; buradan

$$p_A(B) = \frac{m-1}{n-1} < \frac{m}{n} = p(B)$$

ve

$$p_{\bar{A}}(B) = \frac{m}{n-1} > \frac{m}{n} = p(B)$$

yazılır. Eğer A ve B müstakil iseler $p_A(B) = p(B)$ dir.

Bağlı ihtimaller, birinci paragrafta müstakil ihtimaller için yapılana benzer tarzda hesaplanır. N sayıda neticesi, A olayına uygun ve eşit olarak mümkün olan, bir deney gözönüne alalım; bu neticelerden M tanesi B olayına uygun olsun ve $N-M$ tanesi uygun olmasın. Bu halde, $p_A(B) = M/N$ [ve $p_A(\overline{B}) = (N-M/N]$ dir. Meselâ, yukarıda verdiğimiz misâlde $n(n-1)$ sayıda netice eşit olarak mümkündür (ilk çekişte n toptan herhangi birisini, ve ikinci çekişte, geri kalan $n-1$ toptan herhangi birisini, çıkartabiliriz), bunlardan $N = m(n-1)$ tanesi A ya uygundur (ilk çekilişte m tane siyah toptan biri ve ikincide geri kalan $n-1$ toptan biri çıkarılıyor); bu $m(n-1)$

sayıda neticeden $M = m(m-1)$ tanesi B ye uygundur (ilk çekilişte m tane siyah toptan biri, ve ikincide geri kalan $m-1$ siyah toptan biri çıkarılıyor); buradan

$$p_A(B) = \frac{M}{N} = \frac{m(m-1)}{m(n-1)} = \frac{m-1}{n-1}$$

bulunur.

Şimdi, deneyin eşit olarak mümkün olan toplam neticelerinin sayısını K ile gösterelim. Bu K neticeden M tanesi A ve B ye aynı zamanda uygun olduğundan AB nin vuku bulması ihtimali M/K dır. Fakat $M/K = (N/K) \cdot (M/N)$ ve $M/N = p_A(B)$ dir (eşit olarak mümkün olan K neticeden N tanesi A ya uygundur). Şu halde

$$p(AB) = p(A)p_A(B)$$

bulunur. Bu, genel halde iki olayın çarpımının ihtimalini belirlemeyi mümkün kılar; çoğu zaman bu bağıntıya da *bileşik ihtimaller kanunu* denir (paragraf 2 de bileşik ihtimaller kanunu dediğimiz şimdikinin bir özel halidir). Şu halde $p(AB)$ yi bulmak için, A ve B arasında bağlılığı karekterize eden $p_A(B)$ bağlı ihtimalini bilmek lâzımdır, sadece $p(A)$ ve $p(B)$ ihtimallerinin bilinmesi AB olayının ihtimalini belirlemez. B ihtimalinin A nın vuku bulup veya bulmaması ile değişmemesi halinde, yani eğer A ve B müstakil iseler $p_A(B) = p(B)$ dir ve evvelce elde ettiğimiz

$$p(AB) = p(A)p(B)$$

formülü bulunur.

Aşağıdaki özellikler bağlı ihtimallerin tarifinden çıkarılır:

a) $0 \leq p_A(B) \leq 1$; B kesin ise yahut $A \subset B$ ise $p(B) = 1$ dir; B mümkün değilse yahut A ve B uygun değilse $p_A(B) = 0$ dır;

b) eğer $B \subset B_1$ ise $p_A(B) \leq p_A(B_1)$ dir;

c) eğer B ve C uygun değilse $p_A(B + C) = p_A(B) + p_A(C)$ dir; eğer $B_1, B_2, \cdots, B_k$ ikişer ikişer uygun değilse

$$p_A(B_1 + B_2 + \cdots + B_k) = \\ = p_A(B_1) + p_A(B_2) + \cdots + p_A(B_k);$$

dır;

d) $p_A(\overline{B}) = 1 - p_A(B)$ dir.

Bağlı ihtimallerin bu özelliklerinin ispatı, basit (bağlı olmayan) ihtimallerin aynı özelliklerinin ispatına tamamen benzer.

Şuna da işaret edelim ki,

$$p(AB) = p(A)p_A(B)$$

eşitliğinden

$$p(A)p_A(B) = p(B)p_B(A) \quad \text{yahut} \quad \frac{p(A)}{p_B(A)} = \frac{p(B)}{p_A(B)}$$

bulunur (gerçekten, AB ve BA olayları tabiatiyle farklı değildir). Buradan şu netice çıkarılır: A ve B olaylarının $p(A)$ ve $p(B)$ ihtimalleri ve, B olayının A nın vukuuna bağlı olan $p_A(B)$ ihtimali bilindiğine göre $p_B(A)$ ihtimali belirlenebilir:

$$p_B(A) = p_A(B) \cdot \frac{p(A)}{p(B)}$$

Şu halde, yukarıda verilen misâlde $p(A)=p(B)=m/n$ olduğundan (ilk ve ikinci çekilişte bir siyah top çekilmesi ihtimalleri m/n ye eşit olduğundan)

$$p_B(A) = p_A(B) = \frac{m-1}{n-1}$$

dir [$p_B(A)$, ikinci çekilişte bir siyah top çekileceği bilindiğine göre ilk çekilişte bir siyah top çekilmesi ihtimalidir].

Nihayet şuna da işaret edelim ki, A yahut $\bar{A}$ olaylarından biri mutlaka vuku bulacağından, AB («B ve $\bar{A}$») ve $\bar{A}B$ («B ve $\bar{A}$») olaylarının toplamı B olayı ile aynıdır.

$$p(AB)=p(A)p_A(B), \quad p(\bar{A}B)=p(\bar{A})p_{\bar{A}}(B)$$

ve

$$p(AB+\bar{A}B)=p(AB)+p(\bar{A}B)$$

olduğundan (A ve $\bar{A}$ olayları uygun olmadığından) AB ve $\bar{A}B$ olayları da uygun değildir.

$$p(B)=p(A)p_A(B)+p(\bar{A})p_{\bar{A}}(B)$$

bulunur. Daima evvelki misâli gözönünde bulundurarak,

$$p(A)\ m/n, \quad p(\bar{A})=\frac{m-n}{n}, \quad p_A(B)=\frac{m-1}{n-1},$$

$$p_{\bar{A}}(B)=\frac{m}{n-1}$$

ve

$$p(A)p_A(B) + p(\bar{A})p_{\bar{A}}(B) =$$

$$= \frac{m}{n} \cdot \frac{m-1}{n-1} + \frac{n-m}{n} \cdot \frac{m}{n-1} = \frac{m}{n} = p(B)$$

bulunur. Keza, eğer bir α deneyinin ikişer ikişer uygun olmayan $A_1, A_2, \cdots, A_k$ gibi k (ve yalnız k) sayıda neticesi varsa, herhangi bir B olayı $A_1B + A_2B + \ldots + A_kB$ toplamı ile ifade olunabilir ve

$$p(B) = p(A_1)p_{A_1}(B) + p(A_2)p_{A_2}(B) + \cdots + p(A_k)p_{A_k}(B)$$

yazılır. Bu formüle, *toplam ihtimal formülü* denir.

PROBLEM 9. — Üç sandıkta sıra ile,

1) 2 beyaz ve 4 siyah top,
2) 4 beyaz ve 2 siyah top,
3) 3 beyaz ve 3 siyah top

bulunmaktadır. Bu sandıklardan gelişigüzel seçilmiş birisinden bir top çekiliyor. Çekilen topun siyah ve beyaz olması hallerinde seçilen sandığın birinci sandık olması ihtimali nedir?

Bir beyaz top çekilmesi olayı A, bir siyah top çekilmesi olayı $\bar{A}$ olsun; topun birinci sandıktan çekilmesi olayını da B ile gösterelim. Deneyin (üç sandıkta bulunan top sayısına göre) $3 \cdot 6 = 18$ neticesi vardır ve bunlar eşit olarak mümkündür (yani üç sandığın herhangi birisinden bir top çekilmesi eşit olarak mümkündür). Bu 18 neticeden 9 u A ya uygun olup bunlardan da 2 si B ye uygundur. $\bar{A}$ ya uygun 9 netice vardır ve bunların 4 ü B ye uygundur. Buradan

$$p_A(B) = \frac{2}{9} \quad \text{ve} \quad p_{\bar{A}}(B) = \frac{4}{9}$$

bulunur.

PROBLEM 10. — «b a o b a b» kelimesini teşkil eden farfler kartonlar üzerine yazılmıştır. Bu kartonlar karıştırılarak dört karton çekiliyor ve yanyana konuyor. Bu tarzda «baba» kelimesinin elde edilmesi ihtimali nedir?

Birinci harf olarak «*B*» yi çekme olayı, *A*, ikinci harf olarak «*A*» yı çekme olayı *B*, üçüncü harf olarak «*B*» yi çekme olayı *C*, dördüncü harf olarak «*A*» yı çekme olayı *D* olsun; bu şartlar altında aranılan olay *ABCD* yazılır. Bileşik ihtimaller formülünü sıra ile tatbik edersek

$$p(A) = 3/6 = \frac{1}{2}; \quad p(AB) = p(A)p_A(B) = \frac{1}{2} \cdot \frac{2}{5} = \frac{1}{5};$$

$$p(ABC) = p(AB)p_{AB}(C) = \frac{1}{5} \cdot \frac{2}{4} = \frac{1}{10}$$

ve nihayet

$$p(ABCD) = p(ABC)p_{ABC}(D) = \frac{1}{10} \cdot \frac{1}{3} = \frac{1}{30}$$

elde edilir.

PROBLEM 11. — 5 sandık gözönüne alalım, bunlardan ikisinin her birisinde bir beyaz ve 5 siyah top bulunsun; diğer birisinde 2 beyaz 5 siyah top bulunsun; son ikisinin her birinde 3 beyaz ve 5 siyah top olsun. Gelişigüzel seçilmiş bir sandıktan bir top çekiliyor. Bu

topun beyaz olması htimali nedir? Bir beyaz top, iki beyaz top ve üç beyaz top bulunan bir sandıktan bir beyaz top çekme olaylarını sıra ile A_1, A_2, A_3 ile gösterelim;

$$p(A_1)=\frac{2}{5},\ p(A_2)=\frac{1}{5} \quad \text{ve} \quad p(A_3)=\frac{2}{5}$$

dir.

Eğer olay, bir beyaz top çekme olayı ise, toplam ihtimal formülü

$$p(B)=p(A_1)p_{A_1}(B)+p(A_2)p_{A_2}(B)+p(A_3)p_{A_3}(B)$$

$$=\frac{2}{5}\cdot\frac{1}{6}+\frac{1}{5}\cdot\frac{2}{7}+\frac{2}{5}\cdot\frac{3}{8}=\frac{23}{84}$$

dür.

§ 4. Olaylar cebiri ve ihtimalin genel tarifi

Evvelki paragraflarda, A ve B gibi iki olaya, bir üçüncüsünü tekabül ettiren işlemleri tarif ettik; bu işlemlere A ve B olaylarının toplam ve çarpımı adını verdik ve bunları $A+B$ ve AB ile gösterdik. Bu işlemlere bu adları vermemizin sebebi, olayların «toplama» ve «çarpma» kaidelerine benzerliğinden dolayıdır. Böylece, bunların tarifinden, $A+B=B+A$ ve $AB=BA$ olduğu neticesi çıkarılır; keza, $(A+B)C=AC+BC$ bağıntısını da kullandık (bak. sahife 16). Bu paragrafta «olaylar cebiri» ile «sayılar cebiri» arasındaki benzerlikleri ve farkları daha dikkatle inceleyeceğiz.

Aritmetik ve cebirde, tam sayılar, rasyonel sayılar, reel sayılar (rasyonel ve irrasyonel sayılar) ve kompleks

sayılar gibi değişik mahiyette sayılar incelenir. Tekmil hallerde, a ve b gibi iki sayıya diğer iki sayı tekabül ettirilir: bu iki sayının toplamı $a+b$ ve çarpımı ab *dir* *. Üstelik toplama kaideleri çarpma kaidelerine çok benzer; meselâ,

$$a+b=b+a \qquad \text{ve} \qquad ab=ba;$$

$$(a+b)+c=a+(b+c) \qquad \text{ve} \qquad (ab)c=a(bc)$$

dir. Bu benzerlik 0 ve 1 sayılarında da rastlanır, şöyle ki ilki ile toplama ve ikincisi ile çarpma sayıyı değiştirmez:

$$a+0=a \quad \text{ve} \quad a\cdot 1=a.$$

Fakat benzerlik ümit edildiği kadar devam etmez. Meselâ

---

*) Sayılar gözönüne alınmadan, *sayıların toplam ve çarpımının özeliklerine malik olan*, diğer misâllerle de bir toplam ve bir çarpım tarif edilebilir. Meselâ çift sayılar cümlesini ve tek sayılar cümlesini gözönüne alalım (burada ancak iki cümle bahis konusudur!); bu cümleler için iki işlem şu tarzda tarif edilir:

$$0+0=1+1=0, \qquad 0+1=1+0=1$$

(burada 0 çift sayılar cümlesini, 1 tek sayılar cümlesini gösterir; bir çift sayı ile bir tek sayının toplamı bir tek sayıdır), ve

$$0.0=0.1=1.0=0, \qquad 1.1=1$$

(iki çift sayının çarpımı bir çift sayıdır; iki tek sayının çarpımı bir tek saydır). Keza, 3 ile bölme kalanı sıra ile 0 (3 ile bölünebilen sayı), 1 ve 2 olan sayı cümleleri sıra ile 0, 1, 2 ile gösterilirse «üç elemanlı cebir» elde ederiz ve burada toplama ve çarpma

$$0+0=1+2=0, \quad 0+1=2+2=1, \quad 0+2=1+1=2,$$
$$0.0=0.1=0.2=0, \quad 1.1=2.2=1, \quad 1.2=2$$

bağıntıları ile tarif edilir.

$$(a+b)\cdot c=a\cdot c+b\cdot c$$

bağıntısında toplama ve çıkarma simetrik tarzda bulunmazlar. Gerçekten, bu bağıntıda toplama ve çarpma işaretlerinin yerleri değiştirilirse

$$a\cdot b+c=(a+c)\cdot(b+c)$$

bağıntısı bulunur ki abesliği açıkça görülür. Bu sebepten, toplama ve çıkarma işlemlerinin birçok özellikleri çok farklıdır. Meselâ 0 sayısı çarpmada,

$$a\cdot 0=0$$

eşitliği ile ifade olunan çok mühim bir rol oynar (buradan, sıfırdan farklı bir sayının sıfır ile bölünmesi mümkün olmadığı neticesi çıkarılır). Toplamaya,

$$a+1=1$$

benzer eşitliğinin tatbiki yanlıştır.

Bununla beraber toplama ile çarpma arasındaki benzerliğin daha ileri gittiği diğer cebirler mevcuttur. Meselâ, düzlemin bölgeler («şekiller») cümlesini gözönüne alalım. A ve B bölgelerinin $A+B$ toplamı bunların birleşimidir (Şek. 2a). Tabiatiyle

$$A+B=B+A$$

ve

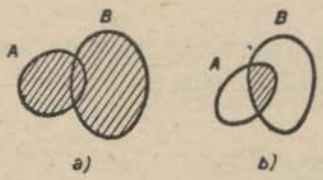
$$(A+B)+C=A+(B+C)$$

dir (bu eşitliğin solunda ve sağında A, B ve C bölgelerinin birleşimi bulunur; bu, daha basit olarak parantez

kaldırılarak, $A + B + C$ şeklinde gösterilir). Sıfır, hiçbir nokta ihtiva etmeyen ve boş bölge denen bir O bölgesi ile gösterilir; bu bölge için

$$A + O = A$$

yazılır.



Şek. 2

A ve B bölgelerinin AB çarpımı, bunların ara kesiti (ortak bölgesi) ile tarif edilmiştir (Şek. 2b).

$$AB = BA$$

ve

$$(AB)C = A(BC)$$

olacağı aşikârdır (bu eşitliğin solunda ve sağında A, B ve C bölgelerinin ortak bölgesi bulunur; bu, daha basit olarak ABC ile gösterilir). Burada birim I düzlemi ile gösterilmiştir. Gerçekten, A bölgesi ne olursa olsun,

$$AI = A$$

dır.

Böylece tarif edilen bölgeler cebirinde

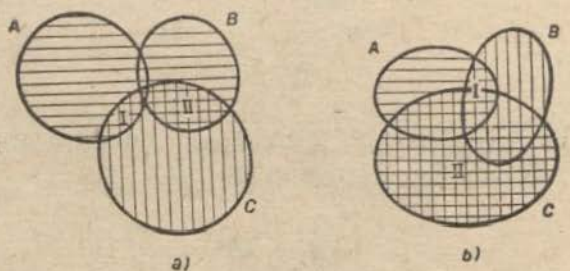
$$(A + B) \cdot C = A \cdot C + B \cdot C$$

bağıntısı câridir; bunu göstermek için şekil 3a yı incele-

mek kâfidir, burada $A + B$ ve C bölgeleri farklı tarzda taranmıştır: bunların $(A + B) \cdot C$ çarpımı (arakesiti) iki farklı tarzda taranmıştır; I ile $A \cdot C$ çarpımı, II ile $B \cdot C$ çarpımı gösterilmiştir. Fakat burada ikinci bir bağıntı daha mevcuttur:

$$A \cdot B + C = (A + B) \cdot (B + C);$$

bu, ilk bağıntıda toplama ve çarpma işaretlerinin yerleri değiştirilmek sureti ile elde edilmiştir.



Şek. 3

Bu bağıntıyı göstermek için Şek. 3b yi incelemek kâfidir, burada $A + C$ ve $B + C$ bölgeleri farklı tarzda taranmıştır; şu halde bunların $(A + C) \cdot (B + C)$ çarpımı iki farklı tarzda taranmış olarak gösterilir; $A \cdot B$ bölgesi I ile, C bölgesi II ile gösterimiştirl.

Bu iki bağıntı arasındaki benzerlik, bölgelerin tekmil toplama ve çarpma kaidelerinin benzerliğini izah eder. Meselâ,

$$A \cdot O = O \quad \text{ve} \quad A + I = I;$$

keza

$$A \cdot A = A \quad \text{ve} \quad A + A = A$$

yazılır; bu eşitliklerin hiç biri sayılar cebiri için muteber değildir. Aritmetik ve cebirde iki sayının mukayesesi mühim bir rol oynar. $\leq$ işaretini kullanarak ($a \leq b$, *a nın b den küçük veya b ye eşit* olduğunu göstermek üzere) iki sayının mukayesesinin temel kaideleri şunlardır:

$a \leq a$ (her a sayısı kendisinden küçük veya kendisine eşittir);

eğer $a \leq b$ ve $b \leq a$ ise $a = b$ dir (eğer a, b den küçük veya b ye eşit ise, ve b, a dan küçük veya a ya eşit ise, a ve b eşittir);

eğer $a \leq b$ ve $b \leq c$ ise $a \leq c$ dir (eğer a, b den küçük veya b ye eşit ise, ve b, c den küçük veya c ye eşit ise, a, c den küçük veya c ye eşittir).

Bölgelerin de mukayesesi yapılabilir: eğer A, B nin bir kısmı ise bu kısım B nin tamamı olabilir ve $A \subset B$ yazılır ($\subset$ işareti $\leq$ işaretinin yerine kullanılır).

Burada da

$$A \subset A,$$

$$\text{eğer } A \subset B \text{ ve } B \subset A \text{ ise } A = B$$

$$\text{eğer } A \subset B \text{ ve } B \subset C \text{ ise } A \subset C$$

olduğu aşikârdır *.

*) Sayıların mukayesesi ile bölgelerin mukayesesi arasındaki esaslı bir farka işaret edelim. a ve b gibi iki (reel) sayı için $a \leq b$ yahut $b \leq a$ bağıntılarından biri daima doğrudur (a ve b sayıları *eşit* ise iki bağıntı da doğru olur). Buna karşılık, A ve B gibi iki bölge için $A \subset B$ ve $B \subset A$ bağıntılarından hiçbiri umumiyetle doğru olmayacaktır. (aynı durum kompleks sayılar için de vakidir; eğer a ve b kompleks sayılarının argümanları aynı ise, ve eğer a nın modülü b nin modülünden küçük veya buna eşit ise, çoğu zaman $a \leq b$ yazılır).

Aşağıdaki kaidelere de işaret edelim:

$$A \subset A + B \quad \text{ve} \quad AB \subset A,$$

keza

$$A \subset I \quad \text{ve} \quad O \subset A$$

(son bağıntı, O boş bölgesinin A bölgesine ait olmayan hiçbir noktayı ihtiva etmediğini gösterir; şunu da ilâve edelim ki A ne olursa olsun, O hiç bir nokta ihtiva etmez).

Bölgeler cebiri ile sayılar cebiri arasındaki esaslı bir fark, her A bölgesine bir A bölgesini (tümlemini) birleştiren bir işlemin mevcut olmasıdır. Bu mütemmim bölge, A bölgesine ait olmayan tekmil noktaları ihtiva eden bölge olarak tarif edilir. Bu yeni işleme ait kaideler şunlardır:

$$A + \overline{A} = I \quad \text{ve} \quad A\,\overline{A} = O$$

$$\overline{O} = I \quad \text{ve} \quad \overline{I} = O$$

$$\overline{\overline{A}} = A$$

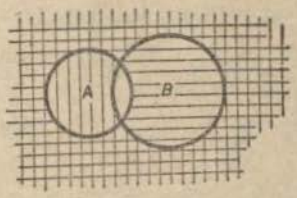
eğer $\overline{A} \subset \overline{B}$ ise $\overline{B} \subset \overline{A}$ dır;

ve nihayet

$$\overline{A + B} = \overline{A}\,\overline{B} \quad \text{ve} \quad \overline{AB} = \overline{A} + \overline{B}$$

(bak. *Şek.* 4, burada $\overline{A}$ ve $\overline{B}$ bölgeleri farklı olarak taranmıştır; iki farklı tarzda taranmış bölge $A + B$ bölgesidir; bir tarzda veya iki farklı tarzda taranmış bölge $\overline{AB}$ bölgesidir).

Bir düzlemin nokta cümlelerinden (domenlerden) gayrı cümleler için de toplanma, çarpma işlemleri $A \subset B$ nev'inden bağıntı ve $\overline{A}$ mütemmim» i tarif edilir. 1 - 3 paragraflarında incelenen tesadüfi olaylar cümlesi buna bir misal teşkil eder; olaylar, kolayca görüleceği gibi, bölgeler cebirinin tekmil özelliklerine maliktir. Tam sayılar



Şek. 4

cümlesi de gözönüne alınabilir; eğer yine iki cümlenin birleşimine ve ara kesitine toplam ve çarpım denirse (meselâ, A_2 ve A_3, sıra ile 2 ve 3 ile bölünebien tam sayılar cümlesi gösterilirse, $A_2 + A_3$ cümlesi tekmil çift sayıları ve tek sayılar arasında 3 ile bölünebilenleri ihtiva edecektir; A_2A_3 cümlesi 6 ile bölünebilen tekmil sayıları ihtiva edecektir), eğer $A \subset B$, işareti A, B nin bir kısmı olduğunu gösteriyorsa (meselâ A_4, 4 ile bölünebilen sayılar cümlesi olmak üzere $A_4 \subset A_2$ dir), eğer $\overline{A}$ ile A ya ait olmayan tam sayılar cümlesi gösterilirse (meselâ A_2, tek sayılar cümlesi ise) ve eğer I ve O sıra ile tekmil tam sayılar cümlesini ve hiçbir sayı ihtiva etmeyen cümleyi gösterirse, yukarıda yazdığımız tekmil bağıntılar muteber olacaktır.

Diğer bir misal bir N tam sayısının bölenler cümlesidir. ($N = 12$ için bu sayılar 1, 2, 3, 4, 6 ve 12 dir); eğer $A + B$ ve AB, sıra ile A ve B nin en küçük ortak

çarpanını ve en büyük ortak bölenini gösterirse, eğer $A \subset B$, «A, B yi böler» manâsına gelirse, eğer O ve I ile 1 ve N (1 den 12 ye kadar) sayıları ve $\overline{A}$ ile N/A $(12/A)$ sayısı gösterilmiş ise, şu bağıntılar da yazılır :

$$A + B = B + A \quad \text{ve} \quad AB = BA$$

$$(A + B) \cdot C = A \cdot C + B \quad \text{ve}$$
$$A \cdot B + C = (A + C) \cdot (B + C)$$

$$A \subset A + B \quad \text{ve} \quad AB \subset A$$

$$\overline{A} + \overline{B} = \overline{A \cdot B} \quad \text{ve} \quad \overline{A \cdot B} = \overline{A} + \overline{B}$$

Nihayet bu tipte çok mühim bir misal, tekmil lojik meseleler cümlesi (yani doğru veya yanlış demenin abes olduğu iddialar) dır; bu cümlenin incelenmesi matematik lojik'in mevzuunu teşkil eder. A ve B gibi iki lojik meselesinin $A + B$ toplamından ve AB çarpımından, sıra ile «ya A, yahut B» ve «A ve B» anlaşılır; $A \subset B$ işareti, A değişiyorsa B de değişiyor (kısaca «A, B yi icap ettirir») demektir. $\overline{A}$, A nın olmadığını gösterir («A meselesi yanlıştır); I ve O, sıra ile bütün dünyaca doğru ve yanlış meseleleri gösterir. Yukarıda yazdığımız tekmil bağıntılar yine muteberdir. Böylece

$$A + \overline{A} = I$$

dış üçtebir kaidesidir: A meselesi daima ya doğrudur ya yanlıştır ;

$$A \cdot \overline{A} = O$$

bağıntısı *mubayenetsizlik kaidesidir:* A meselesi hiçbir halde hem doğru hem yanlış olamaz.

Bütün bu özelliklere malik olan cebirsel sistemlerin ehemmiyeti bu sistemlerin özel olarak incelenmesine yol açar. Bu sistemlere, lojik sahasındaki çalışmaları ile tanınmış ve ilk defa lojik üzerinde araştırmalarında böyle bir cebirden faydalanmış olan XIX. Yüzyılda yaşamış İngiliz matematikçisi George BOOLE'in adına izafeten *Boole cebirleri* denir.

Genel tarzda, Boole cebirinin elemanları sayılar değildir. Fakat çoğu zaman, her A elemanına, aşağıdaki şartları gerçekleyen, bir $|A|$ yahut $p(A)$ sayısı tekabül ettirmesi mümkündür :

$$0 \leq p(A) \leq 1; \qquad p(O) = 0, \quad p(I) = 1;$$

eğer $A \subset B$ ise, $p(A) \leq p(B)$ dir;

eğer $A \cdot B = O$ ise, $p(A + B) = p(A) + p(B)$,

bu sayıya A elemanının *mutlak değeri* yahut norm'u ve Boole cebirinin kendisine de *norme* denir. Misal olarak, kenarı birime eşit olan bir kare içinde düzlem şekiller ailesi söylenebilir (burada kare, Boole cebirinin I elemanının rolünü oynar), burada bir A şeklinin mutlak değeri yahut normu olarak bu şeklin yüzölçüsü, yahut matematik lojikin meseleler cümlesi alınır, burada bir meselenin mutlak değeri (mornu), mesele doğru ise 1, yanlış ise 0 dır. Boole cebirine diğer bir misâl, 1-3 paragraflarında incelenmiş olan olaylar cebiridir; burada bir A olayının mutlak değeri yahut normu $p(A)$ ihtimalidir.

İhtimaller teorisi ile Boole cebiri arasındaki tekabül, bu teorinin konusunun tarifine hizmet edebilir. *İhtimaller teorisi, bir norme Boole cebiri teşkil eden elemanlar cümlesini tetkik eder* denebilir; bu elemanlar, *olaylar* olup bir

A olayının $p(A)$ normu bu olayın *ihtimali*dir. Meselâ, «sandık probleminde (yahut, benzeri her problemde), verilen n eleman («sayı») ile teşkil edilebilen tekmil cümlelerin norme Boole cebiri gözönüne alınabilir; burada (ve aşağıdaki tekmil misallerde) toplam ve çarpım bunların birleşimi ve arakesiti olarak tarif edilir; norma gelince, aşağıdaki şartla belirlenmiştir : yalnız bir elemandan teşkil edilmiş tekmil cümleler için (başka ifade ile, ayrık noktalar için) norm, sürekli olarak $1/n$ ye eşittir. Fakat, yeni görüşümüze göre, aynı BOOLE cebirini norm için daha genel bir şartla gözönüne alabiliriz: ayrık noktalar için normlar, $p_1 + p_2 + \cdots + p_n = 1$ ifadesini gerçekleyen, $p_1, p_2, \cdots, p_n$ pozitif sayılardır (homogen olmayan bir maddeden yapılmış ve tam küp şeklinde olmayan hileli bir zar halinde $n = 6$ ya tekabül eden bu tip bir BOOLE cebiri vardır). Diğer problemleri gözönüne alarak, BOOLE cebirinin elemanları bir AB doğru parçasının tekmil mümkün kısımları olan hale de rastlayabiliriz, burada norm, gözönüne alınan kısmın doğru parçasının bütün uzunluğuna oranıdır; benzer tarzda, bazı defa, elemanları bir düzlem şeklinin veya bir uzay şeklinin tekmil noktaları olan cebirler gözönüne alınır, burada da norm, gözönüne alınan şeklin yüzey veya hacminin şeklin tamamının yüzey veya hacmına oranı olarak tarif edilir (bu nevi problemler için meselâ sahife... de zikredilen yazarların kitabında «geometrik problemler» üzerine problemleri söyleyebiliriz). Tekmil bu hallere de «hileli zar problemi» teşmil edilebilir; başka ifade ile, bir doğru parçasına, bir düzlem şekline yahut verilen üç boyutlu bir şekle ait cümlelerin BOOLE cebirlerine, yukarıda $p(A)$ fonksiyonu için ileri sürülmüş şartlara tabi olan, tamamen keyfi bir norm ithal olunabilir; bu takdirde, ihtimaller teorisi problemleri bakımından gayet zengin yeni bir cümle elde edilecektir.

İhtimaller teorisinin tarifi olarak, biraz yukarıda verdiğimiz tarif alınırsa bu teoriye ait tekmil problemlerde mütekabil BOOLE cebirinin *apriori* verilmesi icap ettiği (başka ifade ile, problemin verilenleri arasında bulunması gerektiği) anlaşılır. Bu takdirde ihtimaller teorisinin esas problemi, bir kompleks olayın ihtimalinin hesabıdır, bu kompleks olay ise, BOOLE cebirinin işlemleri [meselâ $(A+B)^n$ yahut $(A+B+C).(A+D)$] yardımı ile elementer olayların ihtimalleri cinsinden, $A, B, C, D, \ldots$ gibi muayyen elementer olaylardan itibaren teşkil edilmiştir (bu, elementer geometrininkine benzer bir konudur, bu geometride, uzunluklar ve açılar, bilindiği farz olunan diğer uzunluk ve açılar cinsinden hesaplanır; meselâ, bir dik üçgenin hipotenüsünün uzunluğunun dik açı kenarları cinsinden hesaplanması gibi). İhtimaller teorisi bu tarzda tarif edildiği vakit (ilk defa 1917 de S.N. BERNSTEIN tarafından işaret edilmiştir), $p(A)$, $p(B)$, v.s. elementer ihtimallerinin belirlenmesi kalır; teorinin pratik tatbikatının olabilmesi çin, bu ihtimallerin uzun bir deney serisinde başarının yüzdeleri ile intibak eden değerlere malik olmaları icap eder. Paragraf 1 de, bu şartı gerçekleyen bu elemanter ihtimallerin «klasik» belirlenmesini inceledik; bu, «deneyin eşit olarak mümkün olan tam neticeler sistemi» anlamına dayanmaktadır. Bu tam sistemin mevcut olmadığı haller vardır; bu takdirde $p(A)$ yı başka tarzda belirlemek lazımdır (meselâ, A nın bağlı olduğu deney çok sayıda tekrarlanarak $p(A)$ nın bir yaklaşık değeri aranabilir). Fakat elementer ihtimalleri belirlemeyi mümkün kılan metotlar, teorinin esasını teşkil eden, sonraki hesaplara tesir etmezler.

Netice olarak, yukarıda verdiğimiz tekmil misâllerde, BOOLE cebiri muayyen bir «maksimum cebir» in noktalar cümlesinin cebiri olup tesadüfi değildir; BOOLE

cebirinin böyle bir tarifi ihtimaller teorisinin tekmil problemlerinde mümkündür.

Buradan hareket ederek, ihtimaller hesabının temel konusu, tekmil mümkün olayların norme BOOLE cebirinin tetkiki olmayıp, muhtelif kısımları («alt cümleleri») sonradan «olaylar» olarak nazara alınan, muayyen bir «tam elementer olaylar cümlesi» nin tetkiki olarak gözönüne alınabilir. Bu belirlemeyi tamamlamak için, «tam elementer olaylar cümlesi» nin A alt cümlelerine muayyen bir $p(A)$ normu tekabül ettirmek, ve hakikaten bir norme BOOLE cebiri vermek için, alt cümlelerin kendileri ile bunların normunun gerçekleşmesi icap ettiğ şartları («aksiyomları») sıralamak kâfidir. İhtimaller hesabının (1929 da A.-N.KOLMOGOROFF tarafından verilmiş olan) böyle bir aksiometrik inşası, bu paragrafta bazı karışık özelliklerin incelenmesi için faydalandığımız metoda nazaran bazı üstünlüklere maliktir ve bu sebepten bugün daha fazla kullanılmaktadır; fakat bu metodu burada daha etraflı incelemek bizi, konumuzdan hayli uzaklaştırır.

B Ö L Ü M II

ENTROPİ VE İNFORMASYON

§ 1. Belirsizlik derecesinin bir ölçüsü olarak gözönüne alınan entropi

Bu kitabın konusunu teşkil eden, tesadüfî olayların temel özeliği, bunların vukuunda bir kesinliğin bulunmamasıdır; bu kesinsizlik, olayların bağlı oldukları deneylerin neticesinde muayyen bir belirsizlik meydana getirir. Bununla beraber, bahis konusu olan hallere göre belirsizlik derecesinin çok farklı olabileceği tamamen aşikârdır. Eğer deney, ilk rastlanacak karatavuğun rengini belirlemekten ibaret ise bu rengin siyah olacağı hemen hemen kesindir; Zooloji bilginleri beyaz karatavuğun mevcut olabileceğini kabul ediyorlarsa da, çok az kimse böyle bir netice bekleyeceklerdir. İlk rastlanacak olan kimsenin solak olup olmayacağının belirlenmesi bahis konusu ise, burada daha çok belirsizlik olacaktır; burada da neticeyi önceden tereddüt etmeden söylemek mümkündür, fakat, birinci hale nisbeten daha fazla neticeden şüphe edebiliriz. Eğer ilk rastlanacak kimsenin cinsini önceden söylemek bahis konusu olsa netice daha da şüphelidir. Fakat bu deney, meselâ, bizce tamamen bilinmeyen yirmi koşucunun katıldığı bir yarışta birinci gelecek olanı evvelden söylemeye, yahut bir piyango çekilişinde büyük ikramiyeyi kazanacak numarayı belirlemeye nisbeten daha az belirsizdir; ilk rastlanacak kimsenin cinsinin ne olabileceği bir dereceye kadar önceden kestirilebilir, fakat bir kimsenin

ikinci halde ve *bahusus* sonuncu halde bir tahminde bulunması pek mümkün değildir.

Pratikte, bu belirsizlikleri mukayese etmek maksadı ile, bu nevi tekmil deneyler için belirsizlik derecesini hesaplayabilmek önemlidir. Evvelâ *eşit olarak mümkün* k sayıda neticesi olan deneyler halini inceleyeceğiz. Burada belirsizlik derecesinin k ya tabi olduğu aşikârdır: şu halde $k=1$ için neticede belirsizlik yoktur, ve k nın büyük değerleri için, mümkün neticeler çok değişik olduğundan, neticeyi önceden söylemek güçleşir. Şu halde, belirsizlik derecesinin değerlendirilmesi k ya tabi olacaktır. Başka ifade ile, bu, k nın bir fonksiyonu olacaktır. Bu fonksiyon, $k=1$ için sıfıra eşit ve (belirsizlik yok), üstelik, k nın bir artan fonksiyonu olacaktır.

Bu $f(x)$ fonksiyonunu daha tam olarak belirlemek için bunun gerçeklemesi icap ettiği diğer şartları da bulmak lâzımdır. α ve β gibi iki *müstakil* deney gözönüne alalım (birinin neticesi diğerine tesir etmeyen iki deneyin müstakil deneyler olduğunu hatırlatalım). α deneyinin eşit olarak mümkün k sayıda neticesi olduğunu ve β deneyinin de eşit olarak mümkün l sayıda neticesi olduğunu farz edelim. α ve β nın aynı zamanda vuku bulması demek olan $\alpha\beta$ karma deneyini gözönüne alalım. $\alpha\beta$ deneyinin belirsizliğinin, α deneyinkine nisbeten daha büyük olduğu aşikârdır, zira α nın belirsizliğine β nınki ilâve olur. Şu halde $\alpha\beta$ *deneyinin belirsizliği,* α ve β *deneylerine ait belirsizliklerin toplamına eşittir.* $\alpha\beta$ deneyi, (α için mümkün olan k sayıda neticelerin her bir ile birleştirerek elde edilen) $k\,l$ sayıda eşit olarak mümkün neticeye malik olduğundan, $f(k)$ fonksiyonu için aşağıdaki şart bulunur:

$$f(k\,l) = f(k) + f(l).$$

Bu şart bizi; *k sayıda eşit olarak mümkün neticeye malik olan bir deneyin belirsizlik ölçüsü olarak log k sayısını almaya* sevk eder (gerçekten $\log (kl) = \log k + \log l$ dir). Böylece tarif edilmiş belirsizlik ölçüsü hakikaten $k = 1$ için sıfıra eşit olur ve k ile beraber artar *.

Burada logaritma tabanının seçiminin önemi olmadığına işaret edelim, zira malum olan $\log_b k = \log_b a . \log_a k$ formülü mucibince, bir logaritma sisteminden bir diğerine geçmek için $f(k) = \log k$ fonksiyonu bir sabit (sistem dönüştürüm modülü) ile çarpılır, yani belirsizlik ölçüsü için yeni bir birim seçilir. Teknik sahadaki tatbikatta genel olarak, tabanı 2 olan logaritma sistemi seçilir (başka ifade ile, $f(k) = \log_2 k$ alınır). Bu, belirsizlik birimi olarak, eşit olarak mümkün iki neticesi olan bir deneyin (meselâ yazı - tura oyunu) belirsizliğini seçmek demektir; buna *ikili birim* denebilir; teknik alandaki tatbikatta bu birimin kullanılması bölüm IV de daha iyi anlaşılacaktır. İleride, daha çok alışılmış olan ondalık logaritmaları kullanacağız; başka ifade ile, belirsizlik birimi olarak, eşit olarak mümkün 10 neticesi olan bir deneyin belirsizliğini alacağız (bu, içinde numaralanmış 10 top bulunan bir sandıktan bir top çekme deneyi halidir). Bu (*ondalık birim* denebilen) birim, ikili birimden yaklaşık olarak 10/3 defa daha büyüktür (gerçekten, $\log_{2^{10}} 10 \cong 3{,}32 \cong 10/3$ dür).

Eşit olarak mümkün k neticesi olan bir deney için ihtimaller cetveli şudur :

deneyin neticeleri	A_1	A_2	A_3	$\cdots$	A_k
ihtimaller	$1/k$	$1/k$	$1/k$	$\cdots$	$1/k$

*) Logaritmik fonksiyonun, $f(kl) = f(k) + f(l)$, $= f(1) = 0$ ve $k > 1$ için $f(k) > f(l)$ şartlarını gerçekleyen yegâne fonksiyon olduğunu göstermek kolaydır (bak. § 4).

Toplam belirsizlik olarak log k almış bulunuyoruz, şu halde, ihtimali $1/k$ ya eşit olan neticelerin herbiri için belirsizlik $1/k \log k = -1/k \log 1/k$ ya eşittir denebilir. Fakat bu takdirde belirsizlik derecesinin tarifini genişletmek ve ihtimaller cetveli,

deneyin neticeleri	A_1	A_2	A_3
ihtimaller	1/2	1/3	1/6

olan deneyde A_1, A_2, A_3 neticelerinin belirsizlikleri için sıra ile $-1/2 \log 1/2$, $-1/3 \log 1/3$, $-1/6 \log 1/6$ almak tabiidir, bu suretle bu neticenin toplamı belirsizliği

$$-\frac{1}{2}\log\frac{1}{2} - \frac{1}{3}\log\frac{1}{3} - \frac{1}{6}\log\frac{1}{6}$$

ya eşit olur. Benzer tarzda, genel halde, *ihtimaller cetveli*

deneyin neticeleri	A_1	A_2	A_3	...	A_k
ihtimaller	$p(A_1)$	$p(A_2)$	$p(A_3)$	...	$p(A_k)$

olan bir deneyin belirsizlik derecesi olarak

$$-p(A_1)\log p(A_1) - p(A_2)\log p(A_2) -$$
$$- p(A_3)\log p(A_3) - \cdots - p(A_k)\log p(A_k)$$

alınabilir (Bak. bu bölümün 4. paragrafı). İşte bu sayıya α deneyinin *entropi* si adını vereceğiz ve H (α) ile gösterece-ğiz *.

*) Bir deneyin bu entropi ifadesi şu tarzda da izah edilebilir: Herbiri bir $p(A_i)$ ihtimaline malik olan her $A_i (i = 1,2,.., k)$ neti-

Şimdi $H(\alpha)$ ile gösterdiğimiz entropi'nin özelliklerini inceleyelim. A_1 neticesinin $p(A_1)$ ihtimali, meselâ sıfıra eşit ise (A_1 mümkün değil ise), entropinin ifadesinde buna tekabül eden $-p(A_1) \log p(A_1)$ terimi kaldırılacaktır [daha doğrusu bu terimin manası yoktur, zira $\log p(A_1)$ mevcut değildir]. Bu keyfiyet şu tarzda da gerçeklenebilir: $x = 0$ için $x \log x$ ifadesi bir manayı haiz olmadığından, bu ifadenin değeri için $x \log x$ in $x \to 0$ için limitinin değeri alınabilir (kezâ, $x = 90°$ olduğu vakit *tg x cotg* in değeri için $x \to 90°$ için $tg\, x\ cotg\, x$ in limitini almak lazımdır; yalnız bu şartlar altında $tg\, x\, cotg\, x$ in değeri, x ne olursa olsun, 1 e eşittir). Fakat

$$\lim_{x \to 0} (x \log x) = 0$$

olduğunu gerçeklemek kolaydır. Bunu göstermek için, $x = 1/10^n$ olduğuna işaret etmek kâfidir (daima ondalık logaritmaları gözönüne alıyoruz); $n \to \infty$ olduğu vakit $n/10^n$ sıfıra yaklaşır, zira payda paydadan daha çok çabuk artar 10^n nin $n + 1$ rakkamı vardır ve n büyük olduğu vakit bu sayı n den çok daha büyüktür).

---

cesine belirsizlik derecesinin ölçüsü olarak (tekmil neticelerin eşit olarak mümkün ve her birinin ihtimali $p = 1/k$ olduğu, ve belirsizlik derecesinin ölçüsü için $\log k = -\log p$ alındığı hale benzer tarzda) $-\log p(A_1)$ tekabül ettireceğiz. Bu takdirde, α deneyinin belirsizlik derecesinin ölçüsü için, $p(A_1)$ ihtimali ile $-\log p(A_1)$ değerini alan tesadüfî büyüklüğün ortalama değeri alınabilir (Bak. sahife 8).

Burada tarifini yaptığımız entropi anlamı ile fizikteki entropi anlamı arasındaki münasebeti öğrenmek isteyen okuyucu, sahife... zikredilen L. BRILLOUIN'in kitabına müracaat edebilir.

x deneyinin $H(\alpha)$ entropisinin, yalnız, $p(A_1)$, $p(A_2)$, ..., $p(A_k)$ *ihtimallerinden biri birime eşit diğerleri sıfır* olduğu vakit, sıfıra eşit olur. [$p(A_1) + p(A_2) + \cdots + p(A_k) = 1$ olduğunu hatırlatalım]. Bu keyfiyet, x deneyinin belirlirsizlik derecesi olan $H(\alpha)$ nın manası ile bağdaşır; gerçekten, yalnız bu halde, deneyde hiçbir belirsizlik yoktur.

k sayıda mümkün neticeleri olan deneyler arasında, ihtimaller cetveli

deneyin neticeleri	A_2	A_1	A_3	$\cdots$	A_k
ihtimaller	$1/k$	$1/k$	$1/k$	$\cdots$	$1/k$

olan deneyin en büyük belirsizliğe malik olacağı aşikârdır; işte deneyin neticesini önceden kestirmenin en güç olduğu hal budur. Bu tip deneyi α_0 ile göstereceğiz. Bu, α_0 deneyi için maksimum entropiye tekabül eder. $A_1, A_2, \cdots, A_k$ gibi k sayıda neticesi olan bir deney α ise

$$H(\alpha) = -p(A_1)\log p(A_1)\log p(A_2) - \cdots$$

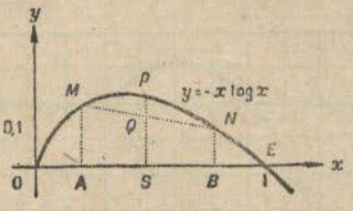
$$\cdots - p(A_k)\log p(A_k) \leq \log k =$$

$$= \underbrace{-\frac{1}{k}\log\frac{1}{k} - \frac{1}{k}\log\frac{1}{k} - \ldots - \frac{1}{k}\log\frac{1}{k}}_{k \text{ defa}} = H(\alpha_0)$$

olur; üstelik eşitlik ancak, $p(A_1) = p(A_2) = \cdots = p(A_k) = 1/k$ ise mümkündür. Şimdilik bu özelliğin tam ispatını yapmıyacağız (Bak. kitabın sonunda ilâve I). Burada teoremi $k = 2$ için ispat etmekle iktifa edeceğiz. Bu halde, şu eşitliği göstermek icap eder:

$$-p(A_1)\log p(A_1) - p(A_2)\log p(A_2) \leq \log 2$$

$x \to 0$ oduğu vaklit $F(x) = -x \log x$ fonksiyonunun sıfıra yaklaştığını biliyoruz; bundan başka, $x = 1$ için de bu fonksiyon sıfır, ve $0 \leq x \leq 1$ için pozitiftir (zira bu halde $\log x$ negatiftir); $x > 1$ için fonksiyon negatif değerler alır. Bu fonksiyonun gösterdiği eğri *şekil* 5 de verilmiştir; $OE = 1$, $OA = p(A_1)$, $OB = p(A_2)$ dir; AM



Şek. 5

ve BN doğru parçaları $-p(A_1)\log p(A_1)$ ve $-p(A_2)\log p(A_2)$ nin değerlerini gösterirler.

$$OA + OB = p(A_1) + p(A_2) = 1$$

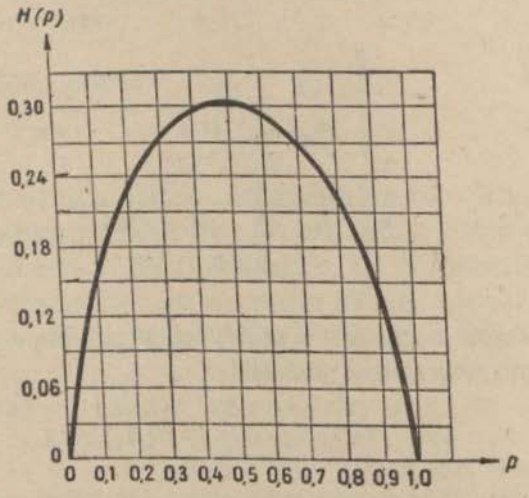
olduğundan, O başlangıç noktasının AB doğru parçasının S orta noktasına uzaklığı 1/2 ye eşittir; şu halde SP doğru parçası $-1/2 \log 1/2 = 1/2 \log 2$ ye eşittir. Fakat AM ve BN doğru parçalarının yarım toplamı, $ABMN$ yamuğunun kenarlarının ortalarını birleştiren SQ doğru parçasına eşit olup bu, SP den küçüktür; buradan

$$1/2\,[-p(A_1)\log p(A_1) - p(A_2)\log p(A_2)] \leq 1/2 \log 2$$

yahut

$$-p(A_1)\log p(A_1) - p(A_2)\log p(A_2) \leq \log 2$$

yazılır, bu ise ancak OA ve OB üstüste geldikleri vakit mümkündür. $H(\alpha)$ entropisini veren formülden şu netice de çıkarılır: $H(\alpha)$ entropisi, $p(A_1), p(A_2), \cdots, p(A_k)$ ihtimallerinin sürekli fonksiyonudur; başka ifade ile, k sayıda mümkün neticeleri olan ve ihtimaller cetveli birbirine yakın olan (yani mütekabil neticelerinin ihtimalleri birbirine yakın olan) α_1 ve α_2 gibi iki deneyin entropi-



Şek. 6

leri de birbirine yakındır. Bu, $F(0) = 0$ şartı ile, $F(x) = -x \log x$ fonksiyonunun sürekli olmasından ileri gelir (*Şek.* 5). Misal olarak *şekil* 6 da, iki mümkün neticesi olan bir deneyin entropisinin, birinci neticenin ihtimali 0 ile 1 arasında değiştiği vakit, değişimini veren $H(p) = -p \log p - (1-p) \log (p-1)$ fonksiyonunun eğrisi verilmiştir.

PROBLEM 12 — Herbirinde 20 top bulunan iki sandık gözönüne alalım, ilkindeki 20 topun 10 u beyaz 5 i si-

yah, 5 i kırmızı, ikincisindeki 20 topun 8 i beyaz 8 i siyah ve 4 ü kırmızı olsun. En az belirsiz olan deney hangisidir?

(α_1 ve α_2 ile göstereceğimiz) bu deneylerin ihtimal cetvelleri şunlardır :

α_1 deneyi (birinci sandıktan bir top çekilmesi) :

çekilen topun rengi	beyaz	siyah	kırmızı
ihtimaller	1/2	1/4	1/4

α_2 deneyi (ikinci sandıktan bir top çekilmesi):

çekilen topun rengi	beyaz	siyah	kırmızı
ihtimaller	2/5	2/5	1/5

Birinci deneyin entropisi

$$H(\alpha_1) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{4}\log\frac{1}{4} - \frac{1}{4}\log\frac{1}{4} \cong$$

$$\cong \frac{1}{2} \cdot 0{,}3010 + \frac{1}{2} \cdot 0{,}6020 = 0{,}4515$$

dir; ikinci deneyin entropisi

$$H(\alpha_2) = -\frac{2}{5}\log\frac{2}{5} - \frac{2}{5}\log\frac{2}{5} - \frac{1}{5}\log\frac{1}{5} \cong$$

$$\cong \frac{4}{5} \cdot 0{,}3979 + \frac{1}{1} \cdot 0{,}6989 = 0{,}4581$$

olup birinciden daha büyüktür.

Şu halde, entropi bir deneyin belirsizlik derecesinin bir ölçüsü olarak gözönüne alınırsa, ikinci deney birinciye nisbeten daha belirsizdir.

Entropi anlamı ilk, defa 1928 de telefon haberleşme mühendisi Amerikalı HARTLEY tarafından düşünülmüş ve k sayıda mümkün derecesi olan bir deneyin belirsizlik derecesini karakterize etmek için $\log k$ sayısını kullanmıştır. HARTLEY, kullanmış olduğu belirsizlik derecesinin ölçüsünün bazı hallere çok uygun düşmekle beraber diğer birçok haller için pek müsait olmadığını anlıyordu; gerçekten, bu tarif, mümkün neticelerin ihtimalleri arasındaki mümkün olan frakları tamamen meçhul bırakır. HARTLEY, değişik mümkün neticeler arasındaki farkların «psikolojik tesirler» den ileri geldiğini, ve şu halde bunların, mühendis ve matematikçiler tarafından değil, psikologlar tarafından nazara alınması icap ettiğini düşünüyordu.

HARTLEY'in hatasını gösteren SHANNON oldu, ve mümkün neticeleri $A_1, A_2, \cdots, A_k$ olan bir α deneyinin belirsizlik derecesinin ölçüsü olarak

$$H(\alpha) = -p(A_1)\log p(A_1) - p(A_2)\log p(A_2) - \cdots$$
$$\cdots - p(A_k)\log p(A_k)$$

büyüklüğünü aldı; burada $p(A_1), p(A_2), \cdots, p(A_k)$, değişik mümkün neticelerin ihtimalleridir; bu büyüklüğe entropi adını veren de SHANNON dur. Başka ifade ile, SHANNON, α deneyinin A_i neticesine $-\log p(A_i)$ ye eşit bir belirsizlik tekabül ettirmiştir (herbirinin ihtimali $p = 1/k$ olan k sayıda eşit olarak mümkün netice halinde, HARTLEY'e göre, belirsizlik olarak $\log k = -\log p$ almak lazımdır); bundan sonra, bütün deneyin belirsizlik derecesi olarak, değişik neticelerin belirsizliklerinin ortalama değeri

alınır [başka ifade ile, bu ihtimaller $p(A_1)$, $p(A_2)$, $\cdots$, $p(A_k)$ olmak *üzere*, $-\log p(A_1)$, $-\log p(A_2)$, $\cdots$, $-\log p(A_k)$ değerlerini alan büyüklüğün ortalama değeridir; 8 inci sahifede verilen ortalama değer tarifi mucibince, $H(\alpha)$ elde edilir]. Böylece, HARTLEY'in esrarlı «psikolojik tesirleri», burada, tamamen matematik (daha doğrusu statistik) anlam olan, ihtimal anlamı ile ithal edimiş olur.

$H(\alpha)$ büyüklüğünün, α deneyinin belirsizlik derecesinin ölçüsü olarak kullanılması birçok hallerde çok pratik fayda sağlar; işte, bu kitabın gayesi ileride bunu göstermektir. Bununla beraber, SHANNON'un, keza HARTLEY'in belirsizlik derecesinin ölçüsü, mümkün olan tekmil anlamlarda «deneyin belirsizliği» ni tâyin eden bütün tesirleri hesaba katamıyacağını unutmamak lâzımdır. Böylece, meselâ $H(\alpha)$ ancak, deneyin değişik mümkün neticelerinin $p(A_1)$, $p(A_2)$, $\cdots$, $p(A_k)$ ihtimallerine tâbidir, fakat bu neticeler birbirlerine (muayyen bir mânada) çok «yakın» veya çok «uzak» olabilmelerine rağmen bunların mahiyetleri ile asla bağlı olmaz. Şu halde «belirsizlik derecesi» ölçümüz, özellikleri aşağıdaki ihtimal cetvelleri ile belirtilmiş olan iki tesadüfî büyüklük için aynı olacaktır:

büyüklüğün değeri	0,9	1	1,1
ihtimal	1/4	1/2	1/4

ve

büyüklüğün değeri	—20	1	1000
ihtimal	1/4	1/2	1/4

bu belirsizlik ölçüsünün derecesi, bir hastanın iki tedavi tarzından birisinde 100 üzerinde 90 halde tam şifa diğer 10 halde hemen hemen şifa ile, diğerinde, 100 üzerinde 90 hal-

de tam şifa diğer 10 halde ölümle neticelenmesinde de aynıdır; böyle iki deney arasındaki esas fark, SHANNON'un entropisinden başka bir büyüklüke karakterize edilmelidir.

$H(\alpha)$ entropisinin bu özelliği, bu büyüklüğün tekmil birçok özellikleri gibi şu tarzda izah edilir: entropi anlamı, evvelâ muayyen haberleşme problemlerinin çözülmesi için ithal olunmuş ve şu halde özellikle bu işe uygulanmıştır. Bir haberin gönderilmesinde geçen zaman, veya bunun fiyatı veyahut gönderilen haberin mânası ile ilgilenilmemesi, $H(\alpha)$ entropisinin, deneyin $A_1, A_2, \cdots, A_i$ neticelerinin mahiyetine tabi olmaması ile izah edilmiştir. Bundan başka, ulaştırma teorisinde muhtelif haberlerin ihtimalleri hiç farklı değildir; mesele etraflı olarak bölüm IV de incelenecektir. Üstelik, büyük sayıda değişik haberler ulaştıran bir ulaştırma hattının çalışmasında, mühim rol oynayan istatistik dağılımdır. Bu sebepten dolayı, ulaştırma hatları ile ilgili meselelerin çözümünde belirsizlik derecesinin ölçüsü, evvelâ, müteakip deney serilerinden teşekkül etmiş muhtelif deneyler için tarif edilmesi icap eder.

Belirsizlik derecesinin bu tarifinde, HARTLEY ve SHANNON'un görüşleri arasındaki farklar ilk bakışta tahmin edilemiyeceği kadar az önemlidir. Filhakika HARTLEY'e göre değişik neticelerin ihtimalleri ile tamamen ilgisiz kalınamaz; aksi halde, k sayıda hakiki neticelere, herhangi sayıda ihtimali sıfıra eşit olan muhayyel neticeler ilâve ederek, deneyimizin mümkün neticelerini sonsuz sayıda artırmak mümkün olacaktı. Şu halde HARTLEY'in tarifinde ihtimali sıfıra eşit olan tekmil «mümkün olmayan neticeleri» i nazara almamak icap eder; fakat, tatbikatta sıfıra eşit kabul edilecek kadar vukuunun ihtimali zaif olan «pratikte imkânsız» neticeleri hesaba katmanın da faydasızlığı düşünülebilir. Şimdi, k

mümkün neicesi olan α deneyi yerine, aynı şartlar altında α deneyini N defa tekrardan ibaret olan α_N deneyini alalım. Bu sonuncu deneyin k^N sayıda mümkün neticesi olacaktır. Bunlar, α deneyinin ilk vukuundaki k sayıda mümkün neticeleri, ikinci vukuundaki k sayıda mümkün neticeleri ve böylece devam ederek, N inci vukuunda k sayıda mümkün neticeleri bir araya getirerek elde edilecektir. Şu halde, HARTLEY'e göre α_N deneyinin belirsizlik derecesi lok $k^N = N \log k$ ya eşittir ve bu suretle bir kere daha α deneyinin belirsizlik derecesinin ölçüsü için log k ifadesini buluruz (gerçekten, α deneyin N defa tekrardan ibaret olan α_N deneyinin belirsizliğinin, α deneyinin belirsizliğinden N defa büyük olacağı tabiidir.

Şimdiye kadar α_N deneyinin k^N sayıda mümkün neticelerinin ihtimalleri için birşey söylemedik. Eğer α deneyinin k sayıda neticelerinin hepsi eşit olarak muhtemel ise, α_N deneyinin k^N sayıda tekmil neticelerinin de eşit olarak muhtemel olacağı aşikârdır; filhakika α_N deneyinin k^N sayıda neticelerinden herhangi birini diğerinden ayırt etmek mümkün değildir. Buna karşılık, α deneyinin k sayıda mümkün neticeleri $p(A_1), p(A_2), \cdots, p(A_k)$ gibi farklı ihtimallere malikse, α_N deneyinin $k = 10^{N \log k}$ sayıda mümkün neticelerinin ihtimalleri de farklı oacaktır. Eğer N çok büyük ile bu $10^{N \log k}$ sayıda neticelerin büyük bir kısmının *tamamen ihmal edilebilir* bir ihtimale malik bulunmaları mümkündür, o derecede ki *tekmil* bu az muhtemel neticelerin ihtimallerinin toplamı çok küçük olur. α_N deneyinin diğer neticelerine (daha muhtemel neticeler) gelince, bunların ihtimalleri, kâfi derecede büyük olmak şartiyle hemen hemen aynı değere malik olabilirler. Daha kesin olarak, *N nin kâfi dercede büyük değerleri için, α_N deneyinin neticelerinin* (genel olarak büyük!) *bir kısmının ihmal edi*lebileceği gösterilebilir, şu tarzda ki ihmâl edilen tekmil

neticelerin ihtimallerinin toplamı, önceden verilen ve istenildiği kadar küçük olan bir sayıdan küçük olsun (bu sayı meselâ, 0,01 yahut 0,001 veyahut 0,000001 alınabilir; bundan başka, bu sayı ne kadar küçük seçilirse N o nisbette büyük alınmalıdır), *ve deneyin ihmâl edilmeyen tekmil neticeleri pratik olarak aynı ihtimale malil bulunsunlar.* Üstelik, ehemmiyetle şu noktaya işaret edelim ki, α^N *deneyinin hemen hemen muhtemel olmayan* neticeleri bu tarzda ihmal edilirse, ihmal edimeyen deneylerin sayısı $10^{N.H(\alpha)}$ mertebesindedir, burada $H(\alpha) = -p(A_1) \log p(A_1) - p(A_2) \log p(A_2) - \ldots - p(A_k) \log p(A_k)$, α deneyinin entropisidir *.

HARTLEY anlamı ile dahi, α^N deneyinin belirsizlik derecesinin ölçüsü olarak $\log 10^{N.H(\alpha)} = N \cdot H(\alpha)$ nın alınması tabiidir (gerçekten, ihtimallerin toplamı çok küçük olan neticeleri ihmal etmek tabiidir); bu takdirde, ilk α deneyinin belirsizlik derecesi olarak $N \cdot H(\alpha)/N$ yani $H(\alpha)$ elde edilir. Şu halde, HARTLEY'in görüşü ile SHANNON'inki arasındaki farkın, SHANNON'un, aynı α deneyinin çok sayıda tekrarı ile elde edilen uzun bir deney silsilesi ithal etmesinde olduğunu görüyoruz; böyle bir deney silsilesinin ithali, ihtimaliyet (veya statistik) bakımdan ziyadesiyle karakteristiktir.

Hemen yukarıda, italik harflerle yazılı özellik, entropi anlamının statistik karakterini gösterir; bu özellik, anla-

---

(*) Buradan şu netice çıkarılır: Özel olarak, α deneyinin tekmil neticeleri eşit olarak muhtemel değilse, ve bunun neticesi $H(\alpha) < \log k$ ise (Bak. sahife 44) α^N deneyinin ihmal edilen neticelerinin sayısı mümkün neticelerinin ziyadesiyle büyük bir kısmını teşkil eder gerçekten, N nin büyük değerleri için

$$10^{N.H(\alpha)}/k^N = 10^{N.H(\alpha)}/10^{N.\log k} = 10^{-N(\log k - H(\alpha))}$$

oranı çok küçüktür.

mın teknik tatbikatının büyük bir kısmının temelini teşkil eder. Fakat bunun ispatı pek basit değildir; bunu, kitabın, hususiyetle haber ulaştırma problemlerine tahsis edilmiş son bölümünde yapacağız.

§ 2. Bağlı entropi

α ve β gibi iki müstakil deney gözönüne alalım, bunların ihtimal cetvelleri şöyle olsun :

α deneyi :

deneyin neticeleri	A_1	A_2	...	A_k
ihtimaller	$p(A_1)$	$p(A_2)$	...	$p(A_k)$

β deneyi:

deneyin neticeleri	B_1	B_2	...	B_l
ihtimaller	$p(B_1)$	$p(B_2)$	...	$p(B_l)$

Aynı zamanda α ve β deneylerinin vukuundan ibaret olan αβ karma deneyini gözönüne alalım. Bu deneyin kj sayıda mümkün neticeleri vardır:

$$A_1B_1,\ A_1B_2,\ \ldots,\ A_1B_l;\ A_2B_1;\ A_2B_2,\ \ldots,\ A_2B_l;\ \ldots$$
$$\ldots;\ A_kB_1,\ A_kB_2,\ \ldots,\ A_kB_l;$$

meselâ A_1B_1, α deneyinin A_1 neticesini ve β deneyinin A_1 neticesini verdiğini gösterir. αβ deneyinin, α ve β deneylerinin her birinden daha belirsiz olduğu aşikârdır, zira birinin belirsizliğine diğerinin belirsizliği ilâve olunur.

(Toplam ihtimaller kanununa benzer olarak *toplam entropiler kanunu* denilen)

$$H(\alpha\beta) = H(\alpha) + H(\beta)$$

bağıntısını gösterelim; bir kere daha işaret edelim ki, bu bağıntı, entropinin, belirsizlik derecesi ölçüsü olarak tefsiri ile bağdaşır.

$$\begin{aligned} H(\alpha\beta) = &- p(A_1B_1)\log p(A_1B_1) - p(A_1B_2)\log p(A_1B_2) \\ &- \cdots - p(A_1B_l)\log p(A_1B_l) - \\ &- p(A_2B_1)\log p(A_2B_1) - p(A_2B_2)\log p(A_2B_2) - \\ &- \cdots - p(A_2B_l)\log p(A_2B_l) - \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &- p(A_kB_1)\log p(A_kB_1) - p(A_kB_2)\log p(A_kB_2) - \\ &- \cdots - p(A_kB_l)\log p(A_kB_l) \end{aligned}$$

yazılır. Fakat

$$p(A_1B_1) = p(A_1)p(B_1),\quad p(A_1B_2) = p(A_1)p(B_2),$$

vs. dir, zira α ve β deneyleri müstakildir (bak. böl. I, § 2, sahife ...). Şu halde denklemin birinci satırının ikinci tarafı şöyle yazılabilir:

$$\begin{aligned} &- p(A_1)p(B_1)\ [\log p(A_1) + \log p(B_1)\ ] - \\ &\quad - p(A_1)p(B_2)\ [\log p(A_1) + \log p(B_2)\ ] - \cdots \\ &\quad\cdots - p(A_1)p(B_l)\ [\log p(A_1) + \log p(B_l)\ ] = \\ &= - p(A_1)\ [p(B_1) + p(B_2) + \cdots + p(B)\ ]\log p(A_1) + \\ &\quad + p(A_1)\ [- p(B_1)\log p(B_1) - p(B_2)\log p(B_2) - \cdots \\ &\quad\cdots - p(B_l)\log p(B_l)\ ] = \\ &\quad = - p(A_1)\log p(A_1) + p(A_1)H(B) \end{aligned}$$

(Gerçekten $p(B_1) + p(B_2) + \cdots + p(B_l) = 1$ dir). Aynı şekilde 2 nci $\cdots$, k ıncı satırlar

$$\begin{aligned} &- p(A_2)\log p(A_2) - p(A_2)H(\beta) \\ &\cdots\cdots\cdots\cdots\cdots \\ &- p(A_k)\log p(A_k) - p(A_k)H(\beta) \end{aligned}$$

yazılır. Nihayet

$$H(\alpha\beta) = -p(A_1)\log p(A_1) - p(A_2)\log p(A_2) - \cdots$$
$$\cdots - p(A_i)\log p(A_i) + [p(A_1) + p(A_2) + \cdots$$
$$\cdots + p(A_i)]\, H(\beta)$$
$$= H(\alpha) + H(\beta)$$

bulunur (gerçekten $p(A_1) + p(A_2) + \cdots + p(A_i) = 1$ dir).

Şimdi α ve β deneylerinin *mustakil olmadıklarını* farz edelim (meselâ α ve β, aynı bir sandıktan arka arkaya iki top çekme deneyleri olabilir). Bu daha genel halde, $\alpha\,\beta$ karma deneyinin entropisinin α ve β entropilerinin toplamına eşit olacağını söyleyemeyiz. Gerçekten, ikinci deneyin neticesi, birinci deneyin neticesi ile *tamamen* belirlenmiş olabilir (α ve β deneyleri, içinde değişik renkte ancak iki top bulunan bir sandıktan arka arkaya iki top çekme deneyleri halinde olduğu gibi); bu halde, $\alpha\beta$ karma deneyinin (bu deneyin belirsizlik derecesinin ölçüsü olan) entropisi, α β nın entropilerinin toplamına değil, sadece α deneyinin entropisine eşit olacağı tabiidir (bunun böyle olduğunu biraz ileride gerçekliyeceğiz). Şu halde $\alpha\beta$ karma deneyinin genel halde entropisini belirleyelim.

H($\alpha\beta$) entropisini, α ve β nın müstakil olduklarını farz etmeden hesaba başlayalım.

$$H(\alpha\beta) = -p(A_1B_1)\log p(A_1B_1) -$$
$$-p(A_1B_2)\log p(A_1B_2) - \cdots$$
$$\cdots - p(A_1B_j)\log p(A_1B_j) -$$
$$-p(A_2B_1)\log p(A_2B_1) - p(A_2B_2)\log p(A_2B_2) - \cdots$$
$$\cdots - p(A_2B_j)\log p(A_2B_j) -$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$-p(A_iB_1)\log p(A_iB_1) - p(A_iB_2)\log p(A_iB_2) - \cdots$$
$$\cdots - p(A_iB_j)\log p(A_iB_j)$$

yazılır, burada da $A_1, A_2, \cdots, A_i$ ve $B_1, B_2, \cdots, B_j$ evvelce olduğu gibi sıra ile α ve β nın mümkün neticelerini gösterirler. Fakat artık, $p(A_iB_1)$, $p(A_iB_2)$, v.s. ihtimalleri yerine basit ihtimal çarpımları koyamayız; zira artık $p(A_iB_1)$ in değeri $p(A_i)p(B_1)$ olmayıp $p(A_i)\,p_{A_i}(B_1)$ dir, burada $p_{A_i}(B_1)$ ile A_i in vukuuna bağlı olan B_1 olayının ihtimali gösterilmiştir. (bak. § 3, böl. I). Bu keyfiyet, ispatın geri kalan kısmına önemli değişiklikler gösterir.

Yukardaki bağıntının birinci satırının ikinci tarafını sadeleştirelim.

$$
\begin{aligned}
&-p(A_i)p_{A_i}(B_1)\,[\log p(A_i) + \log p_{A_i}(B_1)\,] - \\
&\quad -p(A_i)p_{A_i}(B_2)\,[\log p(A_i) + \log p_{A_i}(B_2)\,] - \cdots \\
&\quad \cdots - p(A_i)p_{A_i}(B_j)\,[\log p(A_i) + \log p_{A_i}(B_j)\,] = \\
&= -p(A_i)\,[p_{A_i}(B_1) + p_{A_i}(B_2) + \cdots + \\
&\quad + p_{A_i}(B_j)\,]\log p(A_i) + \\
&\quad + p(A_i)\,[-p_{A_i}(B_1)\log p_{A_i}(B_1) - \\
&\quad - p_{A_i}(B_2)\log p_{A_i}(B_2) - \cdots - p_{A_i}(B_j)\log p_{A_i}(B_j)\,]
\end{aligned}
$$

bulunur. Fakat $B_1 + B_2 + \cdots + B_j$ olayı kesin olduğundan ($B_1, B_2, \cdots, B_j$ neticelerinden en az biri vuku bulur),

$$
\begin{aligned}
p_{A_i}(B_1) + p_{A_i}(B_2) + \cdots + p_{A_i}(B_j) &= \\
= p_{A_i}(B_1 + B_2 + \cdots + B_j) &= 1
\end{aligned}
$$

dir. Diğer taraftan

$$
\begin{aligned}
&p_{A_i}(B_1)\log p_{A_i}(B_1) + p_{A_i}(B_2)\log p_{A_i}(B_2) + \\
&\qquad + \cdots + p_{A_i}(B_j)\log p_{A_i}(B_j)
\end{aligned}
$$

toplamı, A_i olayının gerçekleşmesi şartı ile β deneyinin entropisini gösterir (β nın muhtelif neticelerinin ihtimalleri α nın neticesine bağlı olduğu gibi β nın entropisi

α nın neticesine tabidir). Bu ifadeye β *deneyinin* A_1 *olayının vukuuna bağlı entropisi* denir ve $H_{A_1}(\beta)$ ile gösterilir. Bu tarzda $H(\alpha\beta)$ ifadesinin birinci satırı

$$-p(A_1)\log p(A_1) + p(A_1)H_{A_1}(\beta)$$

yazılır. Aynı tarzda, ikinci, ..., k ıncı satırlar

$$-p(A_2)\log p(A_2) + p(A_2)H_{A_2}(\beta)$$
$$\cdots\cdots\cdots\cdots$$
$$-p(A_k)\log p(A_k) + p(A_k)H_{A_k}(\beta),$$

yazılır, burada $H_{A_1}(\beta), \cdots, H_{A_k}(\beta)$ ile, β deneyinin, A_2, $\cdots$, A_k neticelerinin vukuuna bağlı olan entropileri gösterilmiştir.

Buradan şu formül elde edilir :

$$H(\alpha\beta) = -p(A_1)\log p(A_1) - p(A_2)\log p(A_2) - \cdots$$
$$\cdots - p(A_k)\log p(A_k) + p(A_1)H_{A_1}(\beta) + p(A_2)H_{A_2}(\beta) + \cdots$$
$$\cdots + p(A_k)H_{A_k}(\beta) = H(\alpha) + \{p(A_1)H_{A_1}(\beta) + \cdots$$
$$\cdots + p(A_k)H_{A_k}(\beta)\}.$$

Bu ifadenin birinci terimi α deneyinin entropisini gösterir. İkinci terime gelince, $p(A_1)$, $p(A_2)$, $\cdots$, $p(A_k)$ ihtimalleri ile $H_{A_1}(\beta)$, $H_{A_2}(\beta)$, $\cdots$, $H_{A_k}(\beta)$ değerlerini alan bir tesadüfi büyüklüğün ortalama değerini gösterir. Bu ortalama değere, α *deneyinin vukuuna bağlı* β *deneyinin ortalama entropisi, veya daha basit olarak,* α *nın vukuuna bağlı* β *nın entropisi denir; bunu* $H_\alpha(\beta)$ ile göstererek

$$H_\alpha(\beta) = p(A_1)H_{A_1}(\beta) + p(A_2)H_{A_2}(\beta) + \cdots$$
$$\cdots + p(A_k)H_{A_k}(\beta)$$

yazılır. Nihayet

$$H(\alpha\beta) = H(\alpha) + H\alpha\,(\beta)$$

elde edilir. Bu, bir $\alpha\beta$ karma deneyinin entropisinin genel ifadesidir. Buna, mustakil deneyler haline benzer olarak, *toplam entropiler kanunu* da denebilir.

Şu noktaya işaret edelim ki, bu kitapta incelenecek olan meselelerde temel rolü oynayacak olan bu $H_\alpha(\beta)$ talama entropisidir. Bu,

$$\begin{array}{l} p_{A_1}(B_1),\ p_{A_1}(B_2),\ \cdots,\ p_{A_1}(B_j), \\ p_{A_2}(B_1),\ p_{A_2}(B_2),\ \cdots,\ p_{A_2}(B_j), \\ \cdots\cdots\cdots\cdots\cdots \\ p_{A_k}(B_1),\ p_{A_k}(B_2),\ \cdots,\ p_{A_k}(B_j), \end{array}$$

bağlı ihtimaller cetvelinin 1 inci satır hariç, diğer tekmil satırları bilinmeden A_i neticesinin vukuuna bağlı $H_{A_i}(\beta)$ entropisinin belirenebilmesinden dolayıdır. Şu halde, $H_{A_i}(\beta)$ bağlı entropisi, α nın $k-1$ sayıda neticelerinin fonksiyonu olarak β nın neticelerinin ihtimallerinin değişme tarzına hiçbir suretle tabi değildir; Şu halde bu entropi ancak, α nın, β nın neticeleri üzerindeki çok zaif tesirini gösterir, ve yukardaki cetvel bunun tam bir ifadesidir *. Aksine olarak, α nın neticesinin bilinmesini icap ettirmeyen $H_\alpha(\beta)$ ortalama bağlı entropi, α ve β arasındaki, bağlılığı çok esaslı tarzda ifade eder. Bundan, bu bölümün 3 üncü paragrafında etraflıolarak yeniden bahsedeceeğiz.

$H_\alpha(\beta)$ nın diğer mühim özelliklerini verelim. Bunun pozitif veya sıfır olan bir büyüklük olduğu aşikârdır. Eğer

*) Bu cetvelin (ve α ve β deneylerinin bağlı ihtimallerinin cetvellerinin) bilinmesinin, β deneyinde $B_1, B_2, \ldots, B_j$ neticelerinden biri vuku bulduğu bilindiğine göre, $A_1, A_2, \ldots, A_k$ neticelerinin bağlı ihtimallerinin hesabını mümkün kılacağına işaret edelim (Bak. sahife 25).

tekmil $p(A_1)$, $p(A_2)$, $\cdots$, $p(A_k)$ ihtimalleri sıfırdan farklı ise, yani α deneyinin k sayıda neticesi varsa $H_\alpha(\beta) = 0$ ve eğer yalnız

$$H_{A_2}(\beta) = H_{A_2}(\beta) = \cdots = H_{A_k}(\beta)$$

ise, yani α nın neticesi ne olursa olsun, β nın neticesi tamamen beliri olur. Bu takdirde

$$H(\alpha\beta) = H(\alpha)$$

dir. Eğer α ve β deneyleri birbirine bağlı değilseler $H_{A_1}(\beta) = H_{A_2}(\beta) = \cdots = H_{A_k}(\beta) = H(\beta)$ ve

$$H_\alpha(\beta) = H(\beta)$$

dır. Şu halde $H(\alpha\beta) = H(\alpha) + H_\alpha(\beta)$ formülü şu basit şekli alır: $H(\alpha\beta) = H(\alpha) + H(\beta)$ (Bak. sahife 46).

$H_\alpha(\beta)$ bağlı entropinin daima sıfırla, β deneyinin bağlı olmayan entropisi arasında kaldığına işaret edelim:

$$0 \leq H_\alpha(\beta) \leq H(\beta);$$

bu tarzda, α ve α nın birbirine bağlı olmayarak, β nın neticesinin α nın neticesi ile belirlenmesi halleri hakikaten limit hallerdir.

Bu ifade, entropinin, belirsizlik derecesinin ölçüsü olarak tefsir edilmesiyle uygundur; α deneyinin vukuu β nın belirsizliğini ancak azaltacağı, yahut lüzumunda (α ve β nın birbirine bağlı olmamaları hailnde) bu belirsizliği de-

ğiştiremiyeceği ve fakat hiçbir halde bunu artıramıyacağı aşikârdır *.

Bu özelliğin ispatını kitabın sonunda, ilâve I de yapacağız; burada sadece, α deneyinin ancak A_1 ve A_2 gibi eşit olarak muhtemel iki neticesi olması halini inceleyeceğiz. Bu halde

$$H_\alpha(\beta) = p(A_1)H_{A_1}(\beta) + p(A_2)H_{A_2}(\beta) = \\ = \frac{1}{2}H_{A_1}(\beta) + \frac{1}{2}H_{A_2}(\beta)$$

dır, şöyle ki

$$\frac{1}{2}H_{A_1}(\beta) + \frac{1}{2}H_{A_2}(\beta) \leq H(\beta)$$

eşitsizliğini, yahut

$$\frac{1}{2}[-p_{A_1}(B_1)\log p_{A_1}(B_1) - p_{A_1}(B_2)\log p_{A_1}(B_2) - \cdots$$

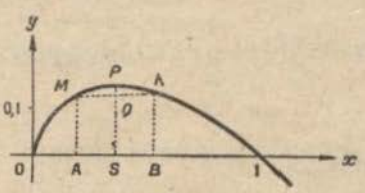
$$\cdots - p_{A_1}(B_j)\log p_{A_1}(B_j)] + \frac{1}{2}[-p_{A_2}(B_1)\log p_{A_2}(B_1) -$$

$$- p_{A_2}(B_2)\log p_{A_2}(B_2) - \cdots - p_{A_2}(B_j)\log p_{A_2}(B_j)] \leq \\ \leq -p(B_1)\log p(B_1) - p(B_2)\log p(B_2) - \cdots \\ \cdots - p(B_j)\log p(B_j),$$

---

*) Tekmil yanlış düşünceleri bertaraf etmek için, $H_{A_i}(\beta)$ bağlı entropinin $H(\beta)$ dan daha büyük veya daha küçük olabileceğine işaret edelim (bak. meselâ ilerideki problem 13 ve 14). Bu keyfiyet, β deneyinin ihtimaller cetvelinin, bir α deneyinin A_i neticesine göre kâfi derecede keyfi tarzda değişebilir (Bak. sahife 21-22).

eşitsizliğini göstermek bahis konusudur, burada B_1, B_2, $\cdots$, B_l ile dairma olduğu gibi, β nın neticeleri gösterilmiştir. $F(x) = -x \log x$ fonksiyonunun gösterdiği eğri üzerinde $OA = p_{A_1}(B_1)$, $OB = p_{A_2}(B_1)$ alalım (Şek. 7); AM



Şek. 7

ve BN doğru parçalarının uzunlukları $-p_{A_1}(B_1) \log p_{A_1}(B_1)$ ve $-p_{A_2}(B_1) \log p_{A_2}(B_1)$ dir

$$-\frac{1}{2} p_{A_1}(B_1) \log p_{A_1}(B_1) - \frac{1}{2} p_{A_2}(B_1) \log p_{A_2}(B_1)$$

toplamı, $ABNM$ yamuğunun kenarlarının orta noktalarını birleştiren SQ doğru parçasının uzunluğuna eşttir. Bundan başka, SP doğru parçası, SQ den büyük olup $-p(B_1) \log p(B_1)$ eşttir, zira

$$OS = \frac{1}{2} \cdot OA + \frac{1}{2} \cdot p(A_1) \cdot p_{A_1}(B_1) + p(A_2) \cdot p_{A_2}(B_1) = p(B_1)$$

dir (bak. toplam ihtimal formülü, sahife 26).

Şu halde

$$-\frac{1}{2} p_{A_1}(B_1) \log p_{A_1}(B_1) - \frac{1}{2} p_{A_2}(B_1) \log p_{A_2}(B_2) \leq \leq -p(B_1) \log p(B_1)$$

bulunur. Keza

$$-\frac{1}{2}p_{A_1}(B_2)\log p_{A_1}(B_2) - \frac{1}{2}p_{A_2}(B_2)\log p_{A_2}(B_2) \leq$$

$$\leq -p(B_2)\log p(B_2)$$

. . . . . .

$$-\frac{1}{2}p_{A_1}(B_j)\log p_{A_1}(B_j) - \frac{1}{2}p_{A_2}(B_j)\log p_{A_2}(B_j) \leq$$

$$\leq -p(B_j)\log p(B_j)$$

eşitsizlikleri gösterilir.

Yukardaki eşitsizlikler toplanarak aranılan eşitsizlik elde edilir.

Şunu da işaret edelim ki, $\alpha\beta$ ve $\beta\alpha$ deneyleri farksız olduğundan $H(\alpha\beta) = H(\beta\alpha)$, yani

$$H(\alpha) + H_\alpha(\beta) = H(\beta) + H_\beta(\alpha)$$

dır. Buradan, α ve β deneylerinin $H(\alpha)$ ve $H(\beta)$ entropileri ve, α nin vukuuna bağlı β deneyinin $H_\alpha(\beta)$ entropisi bilindiği vakit β nın vukuuna bağlı α deneyinin entropisi hesaplanabilir :

$$H_\alpha(\alpha) = H_\beta(\beta) + \{H(\alpha) - H(\beta)\}$$

PROBLEM 13. — m tane siyah ve $n-m$ tane beyaz top ihtiva eden bir sandıktan iki top çekme deneyleri sıra ile α ve β ile gösteriliyor. α ve β deneylerinin $H(\alpha)$ ve $H(\beta)$ entropileri ve aynı deneylerin $H_\alpha(\beta)$ ve $H_\beta(\alpha)$ bağlı entropileri soruluyor.

Tek deneyde bir siyah top veya bir beyaz top çekme olayları A_1 ve A_2, ikinci deneyde aynı olaylar B_1 ve B_2 olsun. Bu iki deney hakkında başka bir şey bilmiyorsak, bu olayların vukuu ihtimalleri şöyle olacaktır:

α deneyi: deney neticeleri	A_1	A_2
ihtimaller	$\frac{m}{n}$	$\frac{n-m}{n}$

β deneyi: deney neticeleri	B_1	B_2
ihtimaller	$\frac{m}{n}$	$\frac{n-m}{n}$

Şu halde bu iki deneyin entropisi aynıdır:

$$H(\alpha) = H(\beta) = -\frac{m}{n}\log\frac{m}{n} - \frac{n-m}{n}\log\frac{n-m}{n}.$$

Eğer α nın neticesini biliyorsak, β nın değişik neticelerinin ihtimalleri artık aynı değere malik olmayacaklardır. Daha açık olarak bunlar

$$p_{A_1}(B_1) = \frac{m-1}{n-1}, \quad p_{A_1}(B_2) = \frac{n-m}{n-1};$$

$$p_{A_2}(B_1) = \frac{m}{n-1}, \quad p_{A_2}(B_2) = \frac{n-m-1}{n-1}$$

olacaktır. Buradan şu netice çıkarılır:

$$H_{A_1}(\beta) = -\frac{m-1}{n-1}\log\frac{m-1}{n-1} - \frac{n-m}{n-1}\log\frac{n-m}{n-1},$$

$$H_{A_2}(\beta) = -\frac{m}{n-1}\log\frac{m}{n-1} - \frac{n-m-1}{n-1}\log\frac{n-m-1}{n-1}$$

$m < n-m$ olması halinde,

$$H_{A_1}(\beta) < H(\beta) \quad , \quad H_{A_2}(\beta) > H(\beta)$$

dir (gerçekten, m tane siyah top ve $m_1 = n - m$ tane beyaz top ihtiva eden bir sandıktan bir top çekmeden ibaret olan bir deneyin belirsizliği, m/m_1 oranının birime yakın olması nisbetinde, büyük olacaktır).

Nihayet,

$$H_\alpha(\beta) = p(A_1)H_{A_1}(\beta) + p(A_2)H_{A_2}(\beta)$$

$$= \frac{m}{n}\left[-\frac{m-1}{n-1}\log\frac{m-1}{n-1} - \frac{n-m}{n-1}\log\frac{n-m}{n-1}\right] +$$

$$\frac{n-m}{n}\left[-\frac{m}{n-1}\log\frac{m}{n-1} - \frac{n-m-1}{n-1}\log\frac{n-m-1}{n-1}\right]$$

ve [$H_\alpha(\beta)$ daima $H(\beta)$ dan küçük olmak üzere],

$$H_\beta(\alpha) - H_\alpha(\beta) + \{H(\alpha) - H(\beta)\} = H_\alpha(\beta)$$

elde edilir.

PROBLEM 14. — Yine m tane siyah ve $n - m$ tane beyaz top ihtiva eden bir sandık gözönüne alınıyor. k tane top çekmeden ibaret olan deney α_k diğer bir top çekmeden ibaret olan deney de β olsun. α_k ve β deneylerinin $H(\alpha_k)$ ve $H(\beta)$ entropileri ve $H_{\alpha_k}(\beta)$, $H_\beta(\alpha_k)$ bağlı entropileri soruluyor ($k = 1$ evvelki problem halidir).

k nın m ve $n - m$ den küçük olduğunu farz edelim. Bu halde, çıkarılan toplar arasında 0, 1, 2, ... yahut k tane siyah top bulunduğuna göre, α_k deneyinin $k + 1$ sayıda neticesi olabilir; bu neticeleri $A_0, A_1, A_2, \ldots, A_k$ ile gösterelim. $A_i (i = 0, 1, 2, \ldots, k)$ neticesinin $p(A_i)$ ihtimali

$$\frac{C_m^i \, C_{n-m}^{k-i}}{C_n^k}$$

oranına eşittir; gerçekten, α_k deneyi için mümkün neticelerin toplam sayısı C_n^k olup (bu n tane toptan teşkil edilebilen k tane top gruplarının sayısıdır) A_i ye uygun neticelerin sayısı ise $C_m^i \cdot C_{n-m}^{k-i}$ dir (m tane siyah top arasından C_m^i tarzda i tane siyah top alınabilir ve $k-i$ tane beyaz top vasıtasiyle C_{n-m}^{k-i} tarzda tamamlanabilir). Buradan anlaşılır ki α_k deneyinin entropisi

$$H(\alpha_k) = -\frac{C_{n-m}^k}{C_n^k} \log \frac{C_{n-m}^{k-1}}{C_n^k} - \frac{C_m^1 \, C_{n-m}^{k-1}}{C_n^k} \log \frac{C_m^1 \, C_{n-m}^{k-1}}{C_n^k} - \frac{C_m^2 \, C_{n-m}^{k-2}}{C_n^k} \log \frac{C_m^2 \, C_{n-m}^{k-2}}{C_n^k} - \ldots$$
$$\ldots - \frac{C_m^{k-1} \, C_{n-m}^1}{C_n^k} \log \frac{C_m^{k-1} \, C_{n-m}^1}{C_n^k} - \frac{C_m^k}{C_n^k} \log \frac{C_m^k}{C_n^k}$$

olacaktır.

β deneyinin iki mümkün neticesi vardır: B_1 bir siyah top çıkarılması ve B_2 bir beyaz top çıkarılması. Bu olayların ihtimalleri m/n ve $(n-m)/n$ dır. Şu halde β deneyinin entropisi

$$H(\beta) = -\frac{m}{n}\log\frac{m}{n} - \frac{n-m}{n}\log\frac{n-m}{n}$$

dir (bak. Problem 13).

Şimdi α_i deneyinin neticesinin A_i olduğunun bilindiğini farz edelim; bu, deneyden sonra sandıkta $m-i$ tane siyah top $n-m-k+i$ tane beyaz top kalıyor demektir; şu halde

$$p_{\alpha_i}(B_1) = \frac{m-i}{n-k}, \qquad p_{\alpha_i}(B_2) = \frac{n-m-k+i}{n-k}$$

ve

$$H_{\alpha_i}(\beta) = -\frac{m-i}{n-k}\log\left(\frac{m-i}{n-k}\right) - \\ -\frac{n-m-k+i}{n-k}\log\left(\frac{n-m-k+i}{n-k}\right)$$

dır. Buradan

$$H_{\alpha_k}(\beta) = \frac{C^k_{n-m}}{C^k_n}\left[-\frac{m}{n-k}\log\left(\frac{m}{n-k}\right) - \right. \\ \left. -\frac{n-m-k}{n-k}\log\left(\frac{n-m-k}{n-k}\right)\right] + \\ + \frac{C^1_m\, C^{k-1}_{n-m}}{C^k_n}\left[-\frac{m-1}{n-k}\log\left(\frac{m-1}{n-k}\right) - \right. \\ \left. -\frac{n-m-k+1}{n-k}\log\left(\frac{n-m-k+1}{n-k}\right)\right] +$$

$$+\frac{C_m^1 C_{n-m}^{k-1}}{C_n^k}\left[-\frac{m-2}{n-k}\log\left(\frac{m-2}{n-k}\right)-\right.$$

$$\left.-\frac{n-m-k+2}{n-k}\log\left(\frac{n-m-k+2}{n-k}\right)+\right.$$

. . . . . .

$$+\frac{C_m^k}{C_n^k}\left[-\frac{m-k}{n-k}\log\left(\frac{n-k}{n-k}\right)-\right.$$

$$\left.\left.-\frac{n-m}{n-m}\log\left(\frac{n-m}{n-k}\right)\right]\right]$$

yazılır. Nihayet, $H_\beta(\alpha)$ bağlı entropi,

$$H_\beta(\alpha_k) = H_{\alpha_k}(\beta) + H(\alpha_k) - H(\beta)$$

formülü vasıtasiyle hesaplanır.

k nın m ve $n-m$ sayılarının birisinden yahut ikisinden de büyük olması halleri yukardakine tamamen benzer tarzda incelenir. Bu hallere tekabül eden formüllerin kurulmasını okuyucuya bırakıp burada bazı hususlara işaret etmekle iktifa edeceğiz.

a) $k \neq n-1$ olduğunu farz edelim. α_{n-1} deneyinin ancak A_1 ve A_2 gibi iki neticesi vardır, bunlardan birincisi sandıkta kalan topun siyah olmasına, ikincisi ise beyaz olmasına tekabül eder. Bu iki neticenin ihtimalleri m/n ve $(n-m)/n$ dir, zira sandıktan çıkarılan $n-1$ topu seçmek demek sandıkta kalan topu seçmek demektir; şu halde α_{n-1} deneyi, n tane top ihtiva eden bir sandıktan

bir top çekmeden ibaret olan α_1 deneyinden farksızdır. Demek ki α_1 deneyinin entropisi

$$H(\alpha_{n-1}) = -\frac{m}{n}\log\frac{m}{n} - \frac{n-m}{n}\log\frac{n-m}{n}$$

ye eşittir, bu ise β deneyinin entropisi ile aynıdır. $H\alpha_{n-1}(\beta)$ bağlı entropiye gelince, bu, aşikâr olarak sıfıra eşittir, zira α_{n-1} deneyinin neticesi, β nın neticesini belirler. Benzer sebeple, $H_\beta(\alpha_{n-1})$ bağlı entropi de sıfıra eşittir.

b) $k = n-2$ olduğunu farz edelim. α_{n-2} deneyinin A_1, A_2 ve A_3 gibi üç mümkün neticesi vardır, bunlar sandıkta iki siyah top, bir siyah ve bir beyaz, yahut iki beyaz top kalması hallerine tekabül eder (burada m ve $n-m$ sayılarının 2 den küçük olmadıklarını farz ediyoruz). Bu neticelerin ihtimalleri

$$p(A_1) = \frac{C_m^2}{C_n^2} = \frac{m(m-1)}{n(n-1)}$$

$$p(A_2) = \frac{C_m^1 \cdot C_{n-m}^1}{C_n^2} = \frac{2m(n-m)}{n(n-1)},$$

$$p(A_3) = \frac{C_{n-m}}{C_n^2} = \frac{(n-m)(n-m-1)}{n(n-1)}.$$

dir. Buradan α_{n-2} deneyinin entropisinin

$$H(\alpha_{n-2}) = -\frac{m(m-1)}{n(n-1)}\log\frac{m(m-1)}{n(n-1)} -$$

$$-\frac{2m(n-m)}{n(n-1)}\log\frac{2m(n-m)}{n(n-1)}-$$

$$-\frac{(n-m)(n-m-1)}{n(n-1)}\log\frac{(n-m)(n-m-1)}{n(n-1)}$$

e eşit olduğu neticesi çıkarılır. β deneyinin, üç neticesinin her birinin vukuuna bağlı entropileri *

$$H_{A_1}(\beta)=0, \quad H_{A_2}(\beta)=\log 2, \quad H_{A_3}(\beta)=0$$

dir. β deneyinin, α_{n-2} nin vukuuna bağlı entropisi

$$H\alpha_{n-2}(\beta)=\frac{2m(n-m)}{n(n-2)}\log 2$$

dir.

Nihayet, α_{n-2} deneyinin, β nın vukuuna bağlı entropisi

$$H(\alpha_{n-2})=H\alpha_{n-2}(\beta)+H(\alpha_{n-2})-H(\beta)$$

$$=\frac{2m(m-n)}{n(n-1)}\log 2-\frac{m(m-1)}{n(n-1)}\log\frac{m(m-1)}{n(n-1)}-$$

$$-\frac{m(n-m)}{n(n-1)}\log\frac{m(n-m)}{n(n-1)}-$$

$$-\frac{(n-m)(n-m-1)}{n(n-1)}\log\frac{(n-m)(n-m-1)}{n(n-1)}+$$

$$+\frac{m}{n}\log\frac{m}{n}+\frac{n-m}{n}\log\frac{n-m}{n}$$

*) Burada $H_{A_3}(\beta)\geq H(\beta)$ dır, zira ancak iki neticesi olan β deneyi log 2 den büyük bir entropiye malik olamaz.

ye eşittir.

c) $m=1$ e eşit ise α_k deneyinin A_1 ve A_2 gibi ancak iki neticesi vardır ve bunlar, tek siyah topun sandıktan çekilen k top arasında bulunması veya sandıkta kalan $n-k$ top arasında bulunması hallerine tekabül eder; bu neticelerin ihtimalleri

$$p(A_1)=\frac{k}{n}, \qquad p(A_2)=\frac{n-k}{n}$$

dır. β deneyinin, α_k nın A_1 neticesinin vukuuna bağlı entropisi sıfıra eşittir:

$$H_{A_1}(\beta)=0,$$

zira A_1 neticesi β nın neticesini tamamen belirler (burada β deneyinde sandıktan çıkarılan top muhakkak beyaz olacaktır). β deneyinin α_k nın A_2 neticesine bağlı entropisi

$$H_{A_2}(\beta)=\frac{1}{n-k}\log\frac{1}{n-k}-\frac{n-k-1}{n-k}\log\frac{n-k-1}{n-k}$$

dır, bu, aynı β deneyinin bağlı olmayan

$$H(\beta)=-\frac{1}{n}\log\frac{1}{n}-\frac{n-1}{n}\log\frac{n-1}{n}$$

entropisinden daha büyüktür (eğer sandıkta bulunan topların biri hariç, hepsi aynı renkte ise, sandıktan bir top çıkarmadan ibaret olan deneyin belirsizlik derecesi, sandığın ihtiva ettiği top sayısının büyüklüğü nispetinde, küçük olacaktır). β deneyinin ortalama entropisine gelince, bu,

$$H_{\alpha_k}(\beta) = \frac{n-k}{n} H_{A_2}(\beta)$$

olup bağlı olmayan $H(\beta)$ entropisinden daha küçüktür.

§ 3. İnformasyon Anlamı

β deneyinin belirsizlik derecesinin mahiyetini tayin eden $H(\beta)$ büyüklüğünü gözönüne alalım, $H(\beta) = 0$, deney neticesinin evvelden bilindiğini gösterir; $H(\beta)$ nın az veya çok büyük olması, deney neticesini az veya çok muhtemel omasını gösterir. β dan evvel yapılan herhangi bir α ölçüsü veya gözlemi, β nın neticelerinin sayısını sınırlıyabilir ve böylece bunun belirsizlik derecesini azaltabilir; meselâ, üç ağırlıktan tartı ile, en ağırını bulmaktan ibaret olan deneyin belirsizlik derecesi, bunlardan herhangi ikisi bir terazide karşılaştırıldıktan sonra, azalır. α ölçüsünün (gözleminin) neticesi β deneyine tesir edebilmesi için, bu neticenin evvelden bilinmemesi tâbiidir; şu halde α, birçok mümkün neticeleri olan, bir yardımcı deney olarak gözönüne alınabilir. α nın vukuu, β nın belirsizlik derecesini azaltması keyfiyeti, β deneyinin α nın vukuuna bağlı $H\alpha(\beta)$ entropisinin, aynı deneyin bağlı olmayan $H(\beta)$ ilk entropisinden küçük olması (daha doğrusu büyük olmaması) ile izah olunur. β deneyi α ya tâbi değilse, α nın vukuu β nın entropisini azaltmaz, yani $H\alpha(\beta) = H(\beta)$ dır; eğer α nın neticesi β nın neticesini tamamen belirlerse, β nın entropisi sıfır olur: $H\alpha(\beta) = 0$. Böylece,

$$I(\alpha, \beta) = H(\beta) - H\alpha(\beta)$$

farkı, α nın vuku bulmasiyle β nın belirsizliğinin ne kadar eksildiğini gösterir; başka ifade ile, bu fark, α

ölçüsü (gözlemi) ile β neticesine ait elde ettiğimiz bilgiyi ölçer; bu farka, β *deneyi hakkında* α *deneyinin bize verdiği informasyon kemmiyeti*, yahut kısaca, β hakkında α nın ihtiva ettiği informasyon denir. Şu halde *informasyonu ölçmek* mümkündür, bu ise birçok hallerde çok faydalıdır. III. ve IV. bölümlerde buna ait misâller vereceğiz.

Entropi ve informasyon anlamları arasındaki münasebet muayyen bir mânada, potansiyel ve potansiyel farkının fizik anlamları arasındaki münasebeti hatırlatır. Entropi, mücerret tarzda «belirsizlik derecesi» ni ölçer; bu anlamın başlıca faydası, bir α deneyinin diğer bir β deneyi üzerindeki tesirini bir $H(\beta) - H\alpha(\beta)$ «entropi farkı» vasıtasiyle ölçmeyi mümkün kılmasıdır. İnformasyon anlamı, β deneyinin şartlarındaki değişikliğe bağlı olduğundan, bu anlam, entropi anlamına nazaran daha «tesirli» dir; şu halde entropi anlamını, informasyon anlamına irca etmek daha faydalıdır ve bu suretle entropinin mânasını daha iyi anlamak mümkün olur β deneyinin $H(\beta)$ entropisi β nın *kendisi hakkında* β *nın ihtiva ettiği informasyon* (gerçekten, β nın vukuu bunun neticesini tamamen belirler, şu halde $H(\beta) = 0$ dır), yahut β hakkında mümkün olan en büyük informasyon (β hakkında «toplam informasyon») olarak tarif edilir. Başka ifade ile, β deneyinin $H(\beta)$ entropisi, bu deneyi icra ederek elde ettiğimiz informasyona, yani β *nın bir neticeisnde bulunan ortalama informasyona* eşittir *. III.

---

*) Entropinin

$$H(\beta) = -p(B_1)\log p(B_1) - p(B_2)\log p(B_2) - \ldots - p(B_j)\log p(B_j)$$

ifadesinin, $p(B_1)$, $p(B_2)$, ..., $p(B_j)$ ihtimalleri ile $-\log p(B_1)$, $-\log p(B_2)$, ..., $-\log p(B_j)$ değerlerini alan, bir ortalama büyüklük manzarası arzettiğine işaret edelim. Şu halde biz B_i neticesinin vukuunun $-\log p(B_i)$ informasyonunu verdiği düşünülebilir. Bu şartlar altında,

ve IV. bölümlerde çok kullanacağımız bu ifadeler, tabiatiyle, «belirsizlik derecesi» ile aynı mânayı haizdirler: bir deneyin belirsizliği ne kadar büyük ise bu deneyin vukuunda elde edilen informasyon o nisbette büyüktür. β hakkında α nın ihtiva ettiği informasyonun, α nın A_i özel neticesine bağlı $H(\beta) - H_{A_i}(\beta)$ büyüklüğünün ortalama değerini gösterdiğini gözönünde bulundurmak lâzımdır; şu halde buna «β hakkında α da bulunan informasyonun ortalama değeri» de denebilir. β hakkında bir informasyon elde etmek için, çoğu zaman, α deneyleri (ölçü, gözlem) arasından seçebilmek mümkün olabilir; meselâ, üç ağırlıktan en ağırını bulmak bahis konusu ise, herhangi bir sırada, ağırlıkları ikişer ikişer karşılaştırmakla işe başlanabilir. Bu takdirde, β hakkında *en fazla informasyon* ihtiva eden α deneyi ile başlamak tercih edilir; diğer bir α deneyi ile başlayarak, *muhtemel olarak,* belirsizlik derecesinde [başka ifade ile, $H(\beta)$ entropisinde] daha az bir eksilme elde ederiz. Hakikatte, α deneyinin, bizim için, α_0 deneyinden daha faydalı olabilmesi müm-

---

meselâ, β deneyinin, ihtimalleri 0,99 ve 0,01 olan B_1 ve B_2 gibi iki mümkün neticesi varsa, B_1 neticesinin vukuunda ancak çok küçük bir informasyon elde edeceğiz: $-\log 0{,}99 \cong 0{,}005$; bu beklenen bir sonuçtur; zira deneyi yapmadan evvel, B_1 neticesinin vuku bulmasına hemen hemen muhakkak nazarı ile bakabiliriz; şu halde bu netice bize pratik olarak β hakkında fazla bir şey öğretmez. Buna karşılık, vuku bulan B_2 neticesi ise elde edilen informasyon $-\log 0{,}01 = 2$, yani birinci hale nazaran çok büyük olacaktır; bu yine beklenen bir sonuçtur; zira deney hakkında daha çok şey öğrenmiş oluyoruz (ümit edilen güç bir olay vuku bulmuştur). Bununla beraber, deneyi çok sayıda tekrar edersek, nâdir olarak bu kadar informasyon elde ederiz; bu sebepten dolayı bir neticede bulunan *informasyon kemmiyetinin ortalama değeri,* tekmil neticelerin aynı ihtimale malik olması haline nisbeten, bu halde daha küçüktür. Şuna da işaret edelim ki, pratik problemlerde, bizi yalnız bu ortalama informasyon kemmiyeti ilgilendirir; biz özel neticeye bağlı informasyon kemmiyeti anlamı, hemen hemen asla kullanılmaz.

kündür; hattâ α_0 deneyinin A neticesi, $H_A(\beta)$ entropisi $H(\beta)$ başlangıç entropisinden büyük olacak kadar, uygun olmayabilir. Bu tabii karşılanmalıdır, zira β nın neticelerinin tesadüfi karakteri, neticesini en çabuk tayin edecek yolu önceden görme imkânına manidir; en fazla yapılabilecek şey en kısa olması en fazla muhtemel olan yolu önceden kestirmektir; işte bu imkân bize informasyon teorisini verir *. $H(\beta) - H_{A_i}(\beta)$ büyüklüklerine gelince, bunlar ancak, β nın mahiyetini belirten büyüklüker olarak nazara alınabilir; gerçekten, eğer α deneyinin A_i neticesini biliyorsak (ve eğer α ve β birbirine bağlı değilseler) β başlangıç deneyini bahis konusu edemeyiz; α nın A_i neticesinin vukuu ile β deneyinin şartlarında hasıl olan değişiklikleri hesaba katmamız icap eder; $H_{A_i}(\beta)$ büyüklüğü sadece yeni bir deneyin entropisi olup A_i neticesi vuku bulması halinde β deneyi buna irca edilir.

PROBLEM 15. — İçinde 5 siyah top ve 10 beyaz top bulunan bir sandıktan bir top çekmeden ibaret olan bir deney gözönüne alınıyor; diğer taraftan, tekrar sandığa koymadan, sandıktan k tane top çekmeden ibaret olan deney α_k olsun. β deneyinin entropisi, ve bu deney hakkında α_1, α_2, α_{13}, ve α_{14} deneylerinin verdiği informasyonlar soruluyor.

β deneyinin entropisi (bak. Problem 12)

---

*) İnformasyon teorisinin, hiç bir halde, β deneyini tamamen belirlemek için lüzumlu, meselâ α deney sayısının bir ortalama değerlendirmesini verebileceği beklememelidir. Meselâ $I(\alpha, \beta)$ informasyonu $H(\beta)$ entropisine eşit ise, α *neticesi ne olursa olsun*, β deneyinin neticesi tamamen belirlenecektir (keza $I(\alpha, \beta)$ informasyonu sıfıra eşitse, α *deneyinin* A_i *neticesi ne olursa olsun*, $H_{A_i}(\beta)$ entropisi, $H(\beta)$ ilk entropisine eşit olacaktır) (Bak. Bölüm III).

$$H(\beta) = -\frac{1}{3}\log\frac{1}{3} - \frac{2}{3}\log\frac{2}{3} \cong 0{,}726$$

dır. 13 ve 14 üncü problemerde elde edilen formülleri kullanarak,

$$I(\alpha_1, \beta) = H(\beta) - H_{\alpha_1}(\beta) = -\frac{1}{3}\log\frac{1}{3} - \frac{2}{3}\log\frac{2}{3} +$$

$$+\frac{1}{3}\left(\frac{2}{7}\log\frac{2}{7} + \frac{5}{7}\log\frac{5}{7}\right) +$$

$$+\frac{2}{3}\left(\frac{5}{14}\log\frac{5}{14} + \frac{9}{14}\log\frac{9}{14}\right) \cong 0{,}002,$$

$$I(\alpha_2, \beta) = H(\beta) - H\alpha_2(\beta) = -\frac{1}{3}\log\frac{1}{3} - \frac{2}{3}\log\frac{2}{3} +$$

$$+\frac{C_{10}^2}{C_{15}^2}\left(\frac{5}{13}\log\frac{5}{13} + \frac{8}{13}\log\frac{8}{13}\right) +$$

$$+\frac{C_{10}^1 \cdot C_5^1}{C_{15}^2}\left(\frac{4}{13}\log\frac{4}{13} + \frac{9}{13}\log\frac{9}{13}\right) +$$

$$+\frac{C_5^2}{C_{15}^2}\left(\frac{3}{13}\log\frac{3}{13} + \frac{10}{13}\log\frac{10}{13}\right) \cong 0{,}0025$$

$$I(\alpha_{13}, \beta) = H(\beta) - H\alpha_{13}(\beta) = -\frac{1}{3}\log\frac{1}{3} - \frac{2}{3}\log\frac{2}{3} -$$

$$-\frac{2 \cdot 5 \cdot 10}{15 \cdot 14}\log 2 \cong 0{,}133$$

ve nihayet

$$I(\alpha_{14}, \beta) = H(\beta) - H_{\alpha_{14}}(\beta) = H(\beta) \ (\cong 0{,}276)$$

elde edilir.

PROBLEM 16. — Muayyen bir N sayısından küçük veya buna eşit olan bir x tam sayısını bulmaktan ibaret olan bir β deneyi ve bu sayının m ile bölünebilir veya böünlemez olduğunu belirtmekten ibaret olan α_m deneyi gözönüne alınıyor; $I(\alpha_m, \beta)$ informasyonu soruluyor.

Aranılan x sayısının 1 ile N arasından bulunan herhangi bir tam sayı olmasının *eşit olarak muhtemel* olduğunu kabul ediyoruz; şu halde β deneyinin eşit olarak muhtemel olan N tane neticesi vardır ve entropisi $\log N$ ye eşittir. α_m deneyinin iki mümkün neticesi vardır: A (aranılan sayı m ile bölünebilir), $\overline{A}$ (aranılan sayı m ile bölünemez); eğer q ile, N nin m ye bölümünden elde edilen sayı gösterilirse, yani, $0 \leq r \leq m$ olmak üzere, $N = qm + r$ ise,

$$p(A) = \frac{q}{N} \qquad \text{ve} \qquad p(\overline{A}) = \frac{N-q}{N}$$

dir (1 ile N arasında bulunan N tane tam sayıdan q tanesi m ile bölünebilir, diğerleri bölünemez).

Şu halde,

$$H_A(\beta) = \log q\,, \qquad H_{\overline{A}}(\beta) = \log (N-q)$$

dür (α_m nin neticesi A ise, β için ancak q tane mümkün hal kalır ve tekmil bunları eşit olarak muhtemel neticeler olarak nazara alıyoruz). Şu halde böylece

$$H\alpha_m(\beta) = \frac{q}{N} \log q + \frac{N-q}{N} \log (N-q)$$

ve

$$I(\alpha_m\beta) = \log N - \frac{q}{N} \log q - \frac{N-q}{N} \log (N-q)$$

elde ederiz.

m büyük olduğu nispette q nün küçük olduğu aşikârdır, ve bunun neticesi olarak, $H_A(\beta)$ ne kadar küçükse $H_{\bar{A}}(\beta)$ o kadar büyüktür. m sayısı ne kadar küçükse, $I(\alpha_m, \beta)$ informasyonu o kadar büyüktür (zira,

$$\frac{q}{N} = \frac{N-q}{N} = \frac{1}{2} \text{ ise}$$

$$I(\alpha_m, \beta) = -\frac{q}{N} \log \frac{q}{N} - \frac{N-q}{N} \log \frac{N-q}{N}$$

maksimum olacaktır *; şu halde x i bulmak için, m nin büyük değeri için muvaffak olmuş bir α_m deneyi α nin belirsizliğini eksiltse bile, $m > 2$ için herhangi bir α_m deneyine tercih edilerek α_2 deneyi ile başlamak lâzımdır.

Şimdi, β deneyi hakkında α deneyinde bulunan informasyonun, α deneyi hakkında β deneyinde bulunan informasyona daima eşit olduğunu gösterelim. Bunun ispatı, evvelki paragrafın neticelerinden hemen çıkarılır; gerçekten,

---

*) $q = 1$ ise, yani $1/2\,N < m \leq N$ ise

$$I(\alpha_m, \beta) = \log N - \frac{N-1}{N} \log (N-1)$$

olacağı aşikârdır; $q = 0$ yani $m > N$ ise $I(\alpha_m, \beta) = 0$ dır.

$$H(\alpha) + H_{\alpha}(\beta) = H(\beta) + H_{\beta}(\alpha)$$

olduğundan

$$I(\alpha, \beta) = H(\beta) - H_{\alpha}(\beta) = H(\alpha) - H\beta(\alpha) = I(\beta, \alpha)$$

dır. Şu halde, β hakkında α deneyinde bulunan informasyona, α *ve* β *deneylerinin birbirlerine ait verilen karşıt informasyon* diyebiliriz.

$$I(\alpha, \beta) = I(\beta, \alpha)$$

formülü

$$I(\alpha, \beta) = H(\alpha) - H_{\beta}(\alpha)$$

da yazılabilir. Buradan şu netice çıkarılır: β deneyi hakkında bir α deneyinde bulunan $I(\alpha, \beta)$ informasyonu, α deneyinin $H(\alpha)$ entropisinden küçük veya buna eşittir; bu, çoğu zaman kullanılan ve aşikâr bir neticedir; zira, bir α deneyinde, bir *diğer* β deneyi hakkında, bulunan informasyon, α da *kendisi hakkında* bulunan informasyondan, yani bu deneyin $H(\alpha)$ entropisinden daha büyük olamaz. Şu halde $H(\alpha)$ entropisi, α *deneyinde bulunan en büyük informasyon* olarak tarif edilebilir (bak. ilerde «toplam informasyon» anlamı).

Aynı $I(\alpha, \beta) = H(\alpha) - H_{\beta}(\alpha)$ formülünden şu netice çıkarılır: ancak $H_{\beta}(\alpha)$ bağlı entropisi sıfıra eşit ise, yani β deneyinin neticesi α yardımcı deneyinin neticesini tamamen belirlerse $I(\alpha, \beta)$ informasyonu α deneyinin $H(\alpha)$ entropisine eşittir; bu, meselâ, problem 16 da böyledir *. $H\beta(\alpha) \neq 0$ ise, $I(\alpha, \beta)$ informasyonu $H(\alpha)$ dan tam $H_{\beta}(\alpha)$ ya eşit büyüklük kadar küçük olacaktır.

---

*) $I(\alpha_m, \beta)$ informasyonu ile $H(\alpha_m)$ entropisinin eşitliğinden, α_m nin A ve $\overline{A}$ neticelerinin ihtimalleri birbirine yakın olduğu nisbette (yani $m = 2$ ise), informasyon o kadar büyük olacaktır.

Şuna işaret edelim ki, $H_\beta(\alpha)$ entropisi sıfıra eşit ise, başka ifade ile, β hakkında α da bulunan $I(\alpha, \beta)$ informasyonu maksimum ise (yani daha fazla informasyon ihtiva eden başka β deneyi mevcut değil ise) β ya tâbi olmayan başka bir γ deneyi hakkında α da bulunan informasyon sıfıra eşittir; bu halde, α deneyi, β deneyinin belirsizliğini kaldırmak için «doğrudan doğruya yönelmiş» tir ve başka hiç bir informasyon ihtiva etmez denebilir.

Bunu ispat etmek için α, β, γ gibi üç deneyin ayrı ayrı vukuundan ibaret olan $\alpha\beta\gamma$ karma deneyini gözönüne alalım: γ, β ya tâbi olmayan *herhangi* bir deneydir. Bu karma deney, $\alpha\beta$ ve γ nın bir kombinezonu, yahut α ve $\beta\gamma$ nın bir kombinezonu olarak nazara alınabilir [*]. Bu takdirde paragraf 2 de incelenen özelliklerin neticesi olarak

$$H(\alpha\beta\gamma) = H(\alpha\beta) + H_{\alpha\beta}(\gamma)$$

ve

$$H(\alpha\beta\gamma) = H(\beta\gamma) + H_{\beta\gamma}(\alpha)$$

yazılır. Diğer taraftan

$$H(\alpha\beta) = H(\beta\alpha) = H(\beta) + H_\beta(\alpha) = H(\beta)$$

dir (zira, hipotezler mucibince $H_\beta(\alpha) = 0$ dır). Bundan başka,

$$H(\beta\gamma) = H(\beta) + H_\beta(\gamma) = H(\beta) + H(\gamma)$$

dır, zira β ve γ deneyleri müstakil farz edilmişlerdir. Nihayet $H_{\beta\gamma}(\alpha)$ bağlı entropisi, $H_\beta(\alpha)$ entropisinden

*) Bu, şu formül ile gösterilebilir:

$$\alpha\beta\gamma = (\alpha\beta)\gamma = \alpha(\beta\gamma).$$

küçük veya buna eşit ve üstelik $H_\beta(\alpha) = 0$ olduğundan $H_{\beta\gamma}(\alpha) = 0$ olur.

Şu halde

$$H(\beta) + H_{\alpha\beta}(\gamma) = H(\beta) + H(\gamma),$$

buradan

$$H_{\alpha\beta}(\gamma) = H(\gamma)$$

bulunur. Diğer taraftan, $H_{\alpha\beta}(\gamma)$ bağlı entropisi $H_\alpha(\gamma)$ bağlı entropisinden küçük veya buna eşittir; şu halde $H_\alpha(\gamma) \geq H(\gamma)$ dır. Halbuki, $H_\alpha(\gamma) \leqq H(\gamma)$ olduğundan

$$H_\alpha(\gamma) = H(\gamma)$$

dır, ve bunun neticesi olarak, ispatı istenilen

$$I(\alpha, \gamma) = H(\gamma) - H_\alpha(\gamma) = 0$$

ifadesi bulunur [*].

§ 4. Özellikleri ile tarif edilen entropi

Bu bölümün 1. paragrafında bahis konusu olan temel kavram, tesadüfi bir neticesi olan bir deneyin entropi yahut belirsizlik derecesi anlamıdır. 1. paragrafta, entropi için verilen tarifin hakikata uygunluğu sezgisel ispatlarla

*) β ya tâbi olmayan bir γ deneyi hakkında bir α deneyinde bulunan informasyonun $H_\beta(\alpha)$ dan küçük veya buna eşit olduğu benzer tarzda gösterilir; $H_\beta(\alpha) = 0$ ise yukarıda gösterilen özellik bulunur.

gerçeklenmiştir; filhakika, 1. bölümün sonunda ve bu bölümde gösterilen teoremlerle bu gerçekleme yapılmış bulunmaktadır. Şimdi bu tarif üzerinde yeniden duracağız, ve entropi anlamının, bir deneyin belirsizlik derecesini ölçmeye mahsus büyüklüğe koşulmuş çok basit bazı şartlardan çıkarıldığını göstereceğiz.

İhtimaler cetveli aşağıda gösterilmiş olan bir α deneyi gözönüne alalım:

deney neticeler	A_1	A_2	$\cdots$	A_k
ihtimaller	$p(A_1)$	$p(A_2)$	$\cdots$	$p(A_k)$

Burada $p(A_1)$, $p(A_2)$, $\cdots$, $p(A_k)$ ihtimallerini p_1, p_2, $\cdots$, p_k ile, ve $H(\alpha)$ entropisini $\Phi(p_1, p_2, \cdots, p_k)$ ile göstereceğiz.

Şimdi, $\Phi(p_1, p_2, \cdots, p_k)$ fonksiyonuna koşulması tabii olan şartları ifade edelim. Evvelâ bu fonksiyon $p_1, p_2, \cdots, p_k$ ihtimallerine tâbi olmalıdır, *bunların sırasına* tâbi değildir; gerçekten, $p_1, p_2, \cdots, p_k$ sayılarının sırasını değiştirmek, ihtimaller cetvelindeki sütunların sırasını değiştirmek demektir, bu ise α deneyini hiçbir suretle değiştirmez. Şu halde ilk şart şudur:

1. *$p_1, p_2, \cdots, p_k$ sayılarının sırası değiştirildiği vakit $\Phi(p_1, p_2, \cdots, p_k)$ fonksiyonu değişmez.*

İkinci şart da şudur:

2. $\Phi(p_1, p_2, \cdots, p_k)$ fonksiyonu süreklidir; başka ifade ile, $p_1, p_2, \cdots, p_k$ ihtimalleri az miktarda değişiktikleri vakit bu fonksiyonda az miktarda değişir; gerçekten, ihtimaller az miktarda değiştikleri vakit, deneyin belirsizlik derecesinin az miktarda değişmesi icap eder.

Üçüncü şart biraz daha karışıktır. Bunun mânasını daha iyi anlamak için, evvelâ, α nın sadece A_1, A_2, A_3 gibi üç mümkün neticesi olduğunu farz edeceğiz; şu halde α nın ihtimaller cetveli şöyledir:

deney neticeleri	A_1	A_2	A_3
ihtimaller	p_1	p_2	p_3

Bu deneyin $H(\alpha)$ belirsizlik derecesi $\Phi(p_1, p_2, p_3)$ dür; bu belirsizlik, deneyin mümkün olan üç neticesinden hangisinin vuku bulacağını bilmediğimizden ileri gelir. O halde iki merhalede, vuku bulan neticeyi belirleyeceğiz. Daha açık ifade ile, evvelâ ilk A_1, A_2 neticelerinden biri mi, yoksa sonuncu A_3 neticesinin mi vuku bulduğunu belirleyeceğiz; bu, α deneyi yerine, ihtimaler cetveli aşağıda gösterilen yeni bir β deneyi almak demektir:

deney neticeleri	B	A_3
ihtimaller	$p_1 + p_2$	p_3

Bu yeni deneyin belirsizlik derecesi $H(\beta) = \Phi(p_1 + p_2, p_3)$ dür. α deneyinin belirsizliği β deneyinkinden büyük olacağı aşikârdır; β deneyinin vukuu α nın neticesini tamamen belirlemez, ve α nın neticesi hakkında muayyen bir belirsizlik kalacaktır; α nın belirsizlik derecesinin β deneyinin belirsizlik derecesini ne kadar aştığını söylemek kolaydır. Eğer α deneyini birçok defa tekrarlarsak ve eğer her defasında evvelâ, β deneyinin B neticesi mi yoksa A_3 neticesi mi vuku bulduğunu belirlersek, bazı hallerde — daha açık ifade ile — neticenin A_3 olduğu tekmil hallerde bu belirleme α nın neticesini belirleyecektir. Neticenin A_1 veya A_2 olması hallerinde, bu neticelerden

hangisinin vuku bulduğunu belirlemek lâzımdır; bu, ihtimaller cetveli aşağıda gösterilmiş olan yeni bir β′ deneyidir:

deney neticeleri	A_1	A_2
ihtimaller	$p_1/(p_1 + p_2)$	$p_2/(p_1 + p_2)$

β′ deneyinin belirsizlik derecesi

$$H(\beta') = \Phi\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right)$$

ye eşittir.

β nın vukuundan sonra β′ deneyinin yapılması icap ettiği hallerin ihtimali (yani ortalama izafi salınım) $p_1 + p_2$ olduğundan, α deneyinin $H(\alpha)$ belirsizliğinin, β dneyinin $H(\beta)$ belirsizlik drecesini $(p_1 + p_2)H(\beta')$ kemmiyeti kadar aştığını nazara almak tabiidir; başka ifade ile,

$$\Phi(p_1, p_2, p_3) = \Phi(p_1 + p_2, p_3) + \\ + (p_1 + p_2)\Phi\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right)$$

eşitliği cari olmalıdır.

İhtimaller cetveli,

deney neticeleri	A_1	A_2	A_3	$\cdots$	A_k
ihtimaller	p_1	p_2	p_3	$\cdots$	p_k

olan bir α deneyi için benzer mülâhazalarla, $\Phi(p_1, p_2, \cdots, p_k)$ fonksiyonunun tâbi olacağı üçüncü şartı şöyle ifade ederiz:

3. $\Phi(p_1, p_2, \cdots, p_k)$ *fonksiyonu*

$$\Phi(p_1, p_2, p_3, \cdots, p_k) =$$
$$= \Phi(p_1+p_2, p_3, \cdots, p_k) + (p_1+p_2)\Phi\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2}\right) \text{(I)}$$

bağıntısını sağlar. Bu bağıntı, ihtimaller cetveli

deney neticeleri	B	A_3	$\cdots$	A_k
ihtimaller	$p_1 + p_2$	p_3	$\cdots$	p_k

olan ve α deneyinin ilk iki neticesini birleştirerek elde edilen β deneyinin $H(\beta)$ belirsizliğinin, α deneyinin $H(\alpha)$ belirsizliğinden $p_1 + p_2$ ile β' deneyinin belirsizliğinin çarpımına eşit bir kemmiyet kadar küçük olduğunu gösterir; β', ilk iki neticeden birinin vukuu halinde hangisinin vuku bulduğunu belirlemekten ibaret olan deneydir.

1, 2 ve 3 şartlarının $\Phi(p_1, p_2, \cdots, p_k)$ fonksiyonunun şeklini tamamen belirleyeceği gösterilebilir: tekmil bu şartları sağlayan yegâne fonksiyon

$$\Phi(p_1, p_2, \cdots, p_k) =$$
$$= c(-p_1 \log p_1 - p_2 \log p_2 - \cdots - p_k \log p_k) \ *$$

şeklindedir.

---

*) Eğer c katsayısına pozitif olma şartı koşulursa, bundan başka, $\Phi(p_1, p_2, \ldots, p_k)$ fonksiyonunun pozitif olmasını açıklamak lâzımdır (tabiatiyle, üç temel şarta bir tamamlayıcı şartın, meselâ $\Phi(1/2, 1/2)$ büyüklüğünün pozitif olması şartını ilâve etmek kâfidir). Şuna da işaret edelim ki, logaritmalar tabanı önceden verilmezse, (*) formülünde c katsayısı kaldırılabilir (gerçekten, $c \log_a p = \log_b p$ dir burada $b = a^{1/c}$dir)

Fakat bu meselenin ispatı haylı güçtür; bununla ilgilenen okuyucu D. - K. FADEFF'in makalesine müracaat edebilir («De l'entropie d'un schéma probabiliste fini», *Progrès des Sciences mathématiques*, t. II, no 1I, Moscou 1956, p. 227-231); biz burada, 1, 2 ve 3 şartlarından çıkarılan ve ispatları çok basitleştiren bir tamamlayıcı şarttan faydalanacağız.

Bundan böyle, *eşit olarak muhtemel k* sayıda neticesi olan α_0 deneyinin belirsizlik ölçüsü $\Phi(1/k, 1/k, \cdots, 1/k)$ fonksiyonu önemli bir rol oynıyacaktır.

α_0 deneyinin tekmil neticeleri eşit olarak muhtemel olduğundan, bu deneyin belirsizlik derecesinin ancak neticelerin sayısına bağlı olacağı aşikârdır:

$$\Phi(1/k, 1/k, \cdots, 1/k) = f(k)$$

olup f fonksiyonu bu bağlılığı ifade eder. α_0 deneyinin belirsizliğinin, muhtemel neticelerin k sayısı büyük odluğu nisbette, büyük olacağı da aşikârdır. Buradan şu şartı ifade edebiliriz:

4. $\Phi(1/k, 1/k, \cdots, 1/k) = f(k)$ *fonksiyonu k nın artan bir fonksiyonudur.*

Şimdi 1, 2, 3 ve 4 şartlarını gerçekleyen $\Phi(p_1, p_2, \cdots, p_k)$ fonksiyonunun mutlaka (*) şeklinde olduğunu gösterelim (burada c, bir pozitif sayıdır). Bunun için 3 şartını ifade eden (I) bağıntısını biraz genelleştirmek lâzımdır.

Evvelâ

$$\Phi(p_1, p_2, \cdots, p_k) = \Phi(p_1+p_2+\cdots+p_i, p_{i+1}, \cdots, p_k) + (p_1+p_2+\cdots+p_i)\Phi\left(\frac{p_1}{p_1+p_2+\ldots+p_i}, \frac{p_2}{p_1+p_2+\ldots+p_i}, \cdots, \frac{p_i}{p_1+p_2+\ldots+p_i}\right)$$

$i > k;$

bağıntısını gösterelim; bu formülün taşıdığı mâna, tabiatiyle, (I) formülününkinin benzeridir; fakat burada α deneyinin i sayıda $A_1, A_2, \cdots, A_i$ neticeleri A gibi tek bir neticede birleştirilmiştir. $i = 2$ için bu formül, (I) formülünün aynı olur: şu halde $i = 2$ için doğrudur. Şimdi i nin herhangi bir değeri için doğru olduğunu kabul edelim; formülü $i = 2$ için kullanarak,

$$\Phi(p_1, p_2, \ldots, p_k) = \Phi(p_1 + p_2 + \ldots + p_i, p_{i+1}, \ldots, p_k) +$$

$$+ (p_1 + p_2 + \ldots + p_i) \cdot \Phi\left(\frac{p_1}{p_1 + \ldots + p_i}, \ldots, \frac{p_i}{p_1 + \ldots + p_i}\right)$$

$$= \Phi(p_1 + p_2 + \ldots + p_i + p_{i+1}, p_{i+2}, \ldots, p_k) +$$

$$+ (p_1 + p_2 + \ldots + p_i + p_{i+1}) \cdot \Phi\left(\frac{p_1 + p_2 + \ldots + p_i}{p_1 + \ldots + p_i + p_{i+1}}, \frac{p_i}{p_1 + \ldots + p_i + p_{i+1}}\right)$$

$$+ (p_1 + p_2 + \ldots + p_i) \cdot \Phi\left(\frac{p_1}{p_1 + \ldots + p_i}, \frac{p_2}{p_1 + \ldots + p_i}, \ldots, \frac{p_i}{p_1 + \ldots + p_i}\right)$$

elde edilir. Diğer taraftan i için eşitliğin doğru olduğunu farz ettiğimizden

$$\Phi\left(\frac{p_1}{p_1 + \ldots + p_i + p_{i+1}}, \ldots, \frac{p_i}{p_1 + \ldots + p_i + p_{i+1}}, \frac{p_{i+1}}{p_1 + \ldots + p_i + p_{i+1}}\right)$$

$$= \Phi\left(\frac{p_1 + p_2 + \ldots + p_i}{p_1 + p_2 + \ldots + p_i + p_{i+1}}, \frac{p_{i+1}}{p_1 + p_2 + \ldots + p_i + p_{i+1}}\right)$$

$$+ \frac{p_1 + p_2 + \ldots + p_i}{p_1 + \ldots + p_2 + p_{i+1}} \cdot \Phi\left(\frac{p_1}{p_1 + p_2 + \ldots + p_i}, \frac{p_2}{p_1 + p_2 + \ldots + p_i}, \ldots, \frac{p_i}{p_1 + p_2 + \ldots + p_i}\right)$$

yazılır; buradan formülün $i + 1$ için de doğru olduğu neticesi çıkarılır:

$$\Phi(p_1, p_2, \cdots, p_k) = \Phi(p_1+p_2+\cdots+p_i,\ p_{i+1}, \cdots, p_k) +$$

$$+ (p_1+p_2+\ldots+p_{i+1}) \cdot \Phi\left(\frac{p_1}{p_1+\ldots+p_{i+1}}, \ldots, \frac{p_i}{p_1+\ldots+p_{i+1}}, \frac{p_{i+1}}{p_1+\ldots+p_{i+1}}\right)$$

Şu halde formül i için de doğrudur.

$\Phi(p_1, p_2, \cdots, p_k)$ fonksiyonu $p_1, p_2, \cdots, p_k$ değişkenlerinin sırasına tâbi olmadığından, yukarıda yaptığımız ispatın neticesi olarak,

$$\Phi(p_1, p_2, \cdots, p_{i+1}, p_i, p_{i+1}, \cdots, p_j, p_{j+1}, \cdots, p_k) =$$

$$= \Phi(p_1, p_2, \cdots, p_{i-1}, p_i + p_{i+1} + \cdots + p_j, p_{j+1}, \cdots, p_k) +$$

$$+ (p_i+p_{i+1}+\cdots+p_j) \cdot \Phi\left(\frac{p_i}{p_i+\ldots+p_j}, \frac{p_{i+1}}{p_i+\ldots+p_j}, \ldots, \frac{p_j}{p_i+\ldots+p_j}\right)$$

$$1 \le i < j \le k,$$

ve genel tarzda

$$\Phi(p_1, \cdots, p_{i_1}, p_{i_1+1}, \cdots, p_{i_2}, p_{i_2+1}, \cdots, p_{i_3}, \cdots, p_{i_s}, \cdots, p_k) =$$

$$= \Phi(p_1+\cdots+p_{i_1}p_{i_1+1}+\cdots+p_{i_2}, \cdots, p_{i_s+1}+\cdots+p_k) +$$

$$+ (p_1 + \cdots + p_{i_1})\Phi\left(\frac{p_1}{p_1+\ldots+p_{i_1}}, \ldots, \frac{p_{i_1}}{p_1+\ldots+p_{i_1}}\right)$$

$$+ (p_{i_s+1} + \cdots + p_{i_2})\Phi\left(\frac{p_{i_2+1}}{p_{i_s+1}+\ldots+p_{i_2}}, \ldots, \frac{p_{i_2}}{p_{i_s+1}+\ldots+p_{i_2}}\right)$$

$$\cdots + (p_{i_s+1}+\cdots+p_k)\Phi\left(\frac{p_{i_s+1}}{p_{i_s+1}+\ldots+p_k}, \ldots, \frac{p_k}{p_{i_s+1}+\ldots+p_k}\right)$$

$$1 \le i_1 < i_2 < \cdots < i_s < k \qquad (2)$$

yazılır. Bu formül, biraz karışık bir şekil altında, biraz genelleştirilmiş toplam entropilerin bir kanunu ifade eder *.

(1) bağlantısının genelleştirilmiş şekli olan (2) formülünün kendisinden ziyade bunun $f(k)$ fonksiyonuna tatbikatı ile ilgileneceğiz. $k = lm$ farz edelim, burada l ve m pozitif tam sayılardır; (2) formülünde bulunan $k = lm$ ihtimallerinin birbirine eşit olduğunu farz edelim (ve bunun neticesi hepsi $1/lm$ ye eşittir); o halde bu eşitliğin sol tarafı $f(lm)$ ye eşit olacaktır. Bundan başka, aynı (2) eşitliğinde bulunan

$$(p_1, \cdots, p_{i_1}), \ (p_{i_1+1} \cdots, p_{i_2}), \ (p_{i_s+1}, \cdots, p_k)$$

gruplarının her birinin l rakkamlarını ihtiva ettiklerini farz edelim; o halde böyle m sayıda grup olacak ve

---

*) Antropiler kanununu (2) formulünden itibaren tesis etmek için

$$i_1 = i_2 - i_1 = i_3 - i_2 = \ldots = k - i_s$$

$$i_1 = i, \ i_2 = 2i, \ i_3 = 3i, \ \ldots, \ k = (s+1)\,i$$

koymak ve $p_1, p_2, \ldots, p_{i_1}; \ p_{i_1+1}, p_{i_1+2}, \ldots, p_{i_2}$ kemmiyetlerini $\alpha\beta$ karma deneyinin $A_1B_1, A_1B_2, \ldots, A_1B_l; \ A_2B_1, A_2B_2, \ldots, A_2B_l; \ldots$ neticelerinin ihtimalleri olarak gözönüne almak kâfidir; bu şartlar altında, $p_1 + p_2 + \ldots + p_{i_1}, \ p_{i_1+1} + p_{i_1+2} + \ldots + p_{i_2}, \ldots$ toplamları, α deneyinin $A_1, A_2, \ldots$ neticelerinin ihtimallerine eşit olacak ve,

$$\frac{p_1}{p_1 + \ldots + p_{i_1}}, \ \frac{p_2}{p_1 + \ldots + p_{i_1}}, \ \ldots \ \frac{p_{i_1}}{p_1 + \ldots + p_{i_1}}$$

$$\frac{p_{i_1+1}}{p_{i_1+1} + \ldots + p_{i_2}} \quad \frac{p_{i_1+2}}{p_{i_1+1} + \ldots + p_{i_2}}, \ \ldots \ \frac{p_{i_2}}{p_{i_1+2} + \ldots + p_{i_2}}$$

oranları da $p_A(B_1), p_A(B_2), \ldots, p_{A_1}(B_l); \ p_{A_2}(B_1), p_A(B_2), \ldots, p_A(B_l); \ldots$ bağlı ihtimallerine eşit olacaktır.

$$p_1 + \cdots + p_{i_1} = p_{i_1+1} + \cdots + p_{i_2} =$$

$$= \cdots = p_{i_{l-1}+1} + \cdots + p_k = l \cdot \frac{1}{lm} = \frac{1}{m};$$

olacaktır; (2) eşitliğinin sağ tarafındaki birinci satır

$$\Phi(1/m, 1/m, \cdots, 1/m) = f(m)$$

olur. (2) eşitliğinin sağ tarafındaki diğer satırlara gelince, bunlar m sayıda olacak ve hepsi

$$(p_1 + \cdots + p_{i_1})\Phi\left(\frac{p_1}{p_1+\cdots+p_{i_1}}, \cdots, \frac{p_{i_1}}{p_1+\cdots+p_{i_1}}\right) =$$

$$= \frac{1}{m}\Phi\left(\frac{1/ml}{1/m}, \cdots, \frac{1/ml}{1/m}\right) = \frac{1}{m}\Phi(1/l, \cdots, 1/l) = \frac{1}{m}f(l)$$

olacaktır. Böylece gözönüne alınan halde (2) denklemi şu basit şekli alır:

$$f(lm) = f(m) + m \cdot \frac{1}{m} f(l) = f(m) + f(l). \quad (2a)$$

Özel olarak,

$$f(k^2) = f(k \cdot k) = f(k) + f(k) = 2f(k),$$
$$f(k^3) = f(k^2 \cdot k) = f(k^2) + f(k) = 3f(k), \cdots, f(k^n) = nf(k)$$

dır.

Daha evvelden $(2a)$ bağıntısının $f(k) = c \log k$ fonksiyonu tarafından sağlandığını biliyoruz. $c \log k$ fonksiyonunun $(2a)$ bağıntısını ve 4. şartı gerçekleyen yegâne fonksiyon olduğunu göstermek kolaydır.

Gerçekten, k ve l herhangi iki pozitif tam sayı olsun. N gibi büyük bir tam sayı alalım ve

$$l^m \leq k^n < l^{m+1}$$

olacak tarzda bir m tam sayısı seçelim. Bu takdirde

$$f(l^m) \leq f(k^n) < f(l^{m+1})$$

yahut

$$mf(l) \leq Nf(k) < (m+1)f(l)$$

olacaktır. Buradan

$$\frac{m}{N} \leq \frac{f(k)}{f(l)} < \frac{m+1}{N}$$

yazılır.

Şimdi

$$l^m \leq k^n < l^{m+1}$$

eşitsizliğinden

$$m \log l \leq N \log k < (m+1) \log l$$

yahut

$$\frac{m}{N} \leq \frac{\log k}{\log l} < \frac{m+1}{N}$$

yazılabileceğine işaret edelim. Şu halde böylece $f(k)/f(l)$ ve $\log k/\log l$ oranlarının birbirine yakın olduklarını görürüz:

$$\left|\frac{f(k)}{f(l)}-\frac{\log k}{\log l}\right|<\frac{1}{N}.$$

Bu son eşitsizlik N sayısı *ne olursa olsun* doğru olduğundan

$$\frac{f(k)}{f(l)}=\frac{\log k}{\log l}$$

yahut

$$\frac{f(k)}{\log k}=\frac{f(l)}{\log l}$$

bulunur. Bu bağıntı her k ve l sayı çifti için doğrudur; şu halde

$$\frac{f(k)}{\log k}=\frac{f(l)}{\log l}=c$$

dir ve c sayısı k ve l ye tâbi değildir; şu halde

$$f(k)=c\log k$$

dir ve $f(k)$ artan fonksiyon olduğuidan $c>0$ dır. Şimdi

$$p_1+p_2+\cdots+p_k=1,$$

$$p_1=\frac{q_1}{p},\ p_2=\frac{q_2}{p},\ \cdots,\ p_k=\frac{q_k}{p}$$

olmak üzere $p_1, p_2, \cdots, p_k$ gibi herhangi rasyonel kesirler gözönüne alalım. (2) formülünü tatbik ederek,

$$f(p) = \Phi\Bigg(\underbrace{\frac{1}{p}, \frac{1}{p}, \cdots, \frac{1}{p}}_{p \text{ defa}}\Bigg) =$$

$$= \Phi\Bigg(\underbrace{\frac{1}{p}, \frac{1}{p}, \cdots, \frac{1}{p}}_{q_1 \text{ defa}}, \underbrace{\frac{1}{p}, \frac{1}{p}, \cdots, \frac{1}{p}}_{q_2 \text{ defa}}, \cdots, \underbrace{\frac{1}{p}, \frac{1}{p}, \cdots, \frac{1}{p}}_{q_k \text{ defa}}\Bigg) =$$

$$= \Phi\left(\frac{q_1}{p}, \frac{q_2}{p}, \cdots, \frac{q_k}{p}\right) + \frac{q_1}{p}\Phi\Bigg(\underbrace{\frac{1}{q_1}, \frac{1}{q_1}, \cdots, \frac{1}{q_1}}_{q_1 \text{ defa}}\Bigg) +$$

$$+ \frac{q_2}{p}\Phi\Bigg(\underbrace{\frac{1}{q_2}, \frac{1}{q_2}, \cdots, \frac{1}{q_2}}_{q_2 \text{ defa}}\Bigg) + \cdots + \frac{q_k}{p}\Phi\Bigg(\underbrace{\frac{1}{q_k}, \frac{1}{q_k}, \cdots, \frac{1}{q_k}}_{q_k \text{ defa}}\Bigg)$$

$$= \Phi(p_1, p_2, \cdots, p_k) + p_1 f(q_1) +$$
$$+ p_2 f(q_2) + \cdots + p_k f(q_k)$$

elde edilir; buradan

$$\Phi(p_1, p_2, \cdots, p_k) = f(p) - p_1 f(q_1) - p_2(q_2) - \cdots - p_k f(q_k)$$
$$= (p_1 + p_2 + \cdots + p_k) f(p) -$$
$$- p_1 f(q_1) - p_2 f(q_2) - \cdots - p_k f(q_k)$$
$$= p_1[f(p) - f(q_1)] + p_2[f(p) - f(q_2)] +$$
$$+ \cdots + p_k[f(p) - f(q_k)]$$

yazılır. Fakat,

$$f(p) - f(q_1) = c \log p - c \log q_1 =$$

$$= -c \log \frac{q_1}{p} = -c \log p_1,$$

$$f(p) - f(q_2) = -c \log p_2, \cdots, f(p) - f(q_k) = -c \log p_k$$

olduğundan kolayca

$$\Phi(p_1, p_2, \cdots, p_k) =$$
$$c(-p_1 \log p_1 - p_2 \log p_2 - \cdots - p_k \log p_k)$$

elde ederiz. Bu formül, $p_1, p_2, \cdots, p_k$ rasyonel sayılar oldukları kabul edilerek ispat edilmiştir. Fakat $\Phi(p_1, p_2, \cdots, p_k)$ fonksiyonu sürekli olduğundan $p_1, p_2, \cdots, p_k$ ne olursa olsun formül doğrudur.

BÖLÜM III

BAZI LOJİK PROBLEMLERİNİN İNFORMASYON ANLAMI YARDIMI İLE ÇÖZÜMÜ

§ 1. Basit Misâller

II. Bölümde incelediğimiz anlam ve meselelerin tatbikatını göstermek amacı ile burada, matematik eğlenceleri hakkında bazı kitaplarda bulunan tipte bir kaç eğlenceli problem tetkik edeceğiz.

PROBLEM 17. — Bir A şehrinin ahalisinin daima hakikatı, A ya yakın bir B şehrinin ahalisinin de daima yalan söyledikleri biliniyor. X ile gösterdiğimiz bir kimse bu iki şehirden birisinde bulunduğunu fakat hangisinde olduğunu bilmiyor. Bu kimse, sualler sorarak hangi şehirde bulunduğunu ve sual sorduğu kimsenin hangi şehirden olduğunu tayin edecektir (A şehrinin ahalisi B ye ve B şehrinin ahalisi A ya gelebilmektedir). X, en az kaç soru ile bu problemlerden birisini veya ikisini çözebilir (X in sual sorduğu kimse ancak «evet» veya «hayır» ile cevap verecektir).

X için, hangi şehirde bulunduğunu belirlemenin bahis konusu olduğunu farz edelim. Burada β deneyinin iki neticesi olabilir (X, A şehrinde yahut B şehrinde bulunabilir). X, hangi şehirde bulunduğuna dair önceden hiç bir fikre sahip olmadığı kabul edilirse, şu iki neticeyi eşit

olarak muhtemel nazara almak lâzımdır. β deneyinin $H(\beta)$ entropisi (yani β nın kendisinde bulunan ve deney vuku bulduğu vakit X tarafından elde edilen informasyon miktarı) $\log 2$ ye ezittir. Fakat, rastlanan kimseye bir soru sormaktan ibaret olan α deneyinin de ancak iki mümkün neticesi vardır (cevap müspet veya menfi olabilir); bu deneyin (cevabın ihtiva ettiği informasyon miktarına eşit olan) $H(\alpha)$ entropisi en fazla $\log 2$ ye eşittir. Problemde bahis konusu olan, β hakkında α deneyinde bulunan $I(\alpha, \beta)$ informasyonunu, β nın $H(\beta) = \log 2$ entropisine eşit olacak tarzda, α sorusunu sormanın, başka ifade ile, β nın neticesini tamamen belirleyecek tarzda β yı bulmanın mümkün olup olmadığıdır. $I(\alpha, \beta)$ ve $H(\beta)$,

$$I(\alpha, \beta) \leq H(\alpha) \quad [I(\alpha, \beta) = H(\alpha) - H_{\beta}(\alpha)]$$

ile bağlı olduklarından ve $H(\alpha)$ entropisi $\log 2$ ye eşit olabileceğinden, burada

$$I(\alpha, \beta) = H(\beta)$$

olduğu ümit edilebilir. Bunun için α sorusu, iki mümkün cevap eşit olarak muhtemel olacak tarzda olmalı *) [$H(\alpha) = \log 2 = H(\beta)$ eşitliği ancak bu şartlar altında mümkündür], ve β nın neticesi α nın neticesini tamamen belirlemelidir [başka ifade ile, $H_{\beta}(\alpha) = 0$ olmalıdır; bu takdirde α sorusu, β nın belirsizliğini kaldırmak için «doğrudan doğruya yönelmiş» olup hiç bir yabancı informasyon ihtiva etmez]. «Bu şehirden misiniz» sorusu tekmil

*) Burada X in A veya B de bulunmasını ve soru sorulan kimsenin A veya B den olmasını eşit olarak muhtemel kabul etmek lâzımdır.

bu şartlara cevap verir (müspet bir cevap *A* şehrinde, menfî bir cevap ancak *B* şehrinde verilebilir).

X in tek bir soru sorduğu kimsenin hangi şehirden olduğunu tayin edebielceğine de işaret edelim (bunun için, cevabını bildiği herhangi bir soru sorması kâfidir, meselâ, «bu şehirde mi bulunuyorum?» yahut «iki kerre iki dört mü eder?»).

Eğer *X* aynı zamanda hangi şehirde bulunduğunu ve soru sorduğu kimsenin hangi şehirden olduğunu belirlemek istiyorsa, β_1, *X* in hangi şehirde bulunduğunu belirleme ve β_2, soru sorulan kimsenin hangi şehirden olduğunu belirleme problemi olmak üzere, $\beta_1\beta_2$ karma broblemini çözmek lâzımdır. Bu deneyin $H(\beta_1\beta_2)$ entropisi β_1 deneyinin $H(\beta_1)$ entropisinden büyüktür zira

$$H(\beta_1\beta_2) = H(\beta_1) + H_{\beta_1}(\beta_2)$$

dir (§ 2, bölüm II.). Başka ifade ile, log 2 den büyük bir informasyon elde etmemiz icap eder [zira $H(\beta_1) = \log 2$ dir]. Bir soru sormaktan ibaret olan α deneyinin $H(\alpha)$ entropisi log 2 den büyük olamıyacağından [ilgi çekici olmayan $H_{\beta_1}(\beta_2)$ bağlı entropinin sıfıra eşit olması halinde, yani *A* şehrinden olanların *B* ye ve *B* şehrinden olanların *A* ya gelememesi halinin dışında), tek bir soru bize bu informasyonu veremez ve $\beta_1\beta_2$ nin neticesini tamamen belirlemeyi mümkün kılmaz. Böylece, informasyon teorisi, *X* in, hangi şehirde bulunduğunu ve soru sorulan kimsenin hangi şehirden olduğunu aynı anda belirlemenin mümkün olmadığını açık olarak ispat eder. Eğer *X* iki soru sorarsa (yani, dört neticesi olan $\beta_1\beta_2$ karma deneyini yaparsa), tabiatiyle $\beta_1\beta_2$ deneyinin neticesini tayin edebilir (bir β_1 sorusu ile β_1 in neticesini, sonra bir β_2 sorusu ile β_2 nin neticesini tayin emek kâfidir).

Şimdi biraz daha karışık şartlarla 17 problemini yeniden ele alalım:

PROBLEM 18. — $A, B, C,$ üç şehir olsun; A nın ahalisinin daima doğru söylediğini, B ninkilerin daima yalan söylediğini, C ninkilerin ise yüzde elli doğru, yüzde elli yalan söylediğini kabul edelim, X, hangi şehirde bulunduğunu ve hangi şehirin ahalisi ile konuştuğunu tayin etmek istiyor. Soru sorduğu kimsenin ancak «evet» veya «hayır» ile cevap verdiği bilindiğine göre en az kaç soru sormalıdır?

Burada, β nın mümkün olan dokuz neticesinden hangisinin vuku bulduğunu bilmek bahis konusudur (X in, A, B ve C şehirlerinden herhangi birinde bulunabileceği ve soru sorduğu kimsenin bu üç şehrin birisinden olabileceğine göre dokuz netice olacağı aşikârdır). Eğer X in önceden β hakkında hiç bir informasyona sahip olmadığı kabul edilirse, bu dokuz neticeyi eşit olarak muhtemel kabul etmek lâzımdır; şu halde β nın $H(\beta)$ entropisi (ve bunun neticesi olarak lüzumlu informasyon miktarı) log 9 'a eşittir. k tane soru sormaktan ibaret olan karma deney $A_k = \alpha_1\alpha_2, \cdots, \alpha_k$ olsun. Her soru için müspet veya menfi bir cevap olabileceğinden, $\alpha_1, \alpha_2, \cdots, \alpha_k$ deneylerinin her birisinin entropisi log 2 ye eşit veya bundan küçük olacaktır. Bundan başka

$$H(\alpha_1\alpha_2) = H(\alpha_1) + H_{\alpha_1}(\alpha_2) \leq H(\alpha_1) + H(\alpha_2)$$

dır [zira $H_{\alpha_1}(\alpha_2) \leq H(\alpha_2)$ dir], genel olarak,

$$H(A_k) = H(\alpha_1\alpha_2, \cdots, \alpha_k) \leq H(\alpha_1) + \\ + H(\alpha_2) + \cdots + H(\alpha_k) \leq k \log 2$$

dir (bu eşitsizliği rekürans yolu ile ispat etmek kolaydır). Bu, şöyle de ifade olunabilir: soruların herbiri log 2 den büyük bir informasyon veremiyeceğinden, k soru sormakla k log 2 den büyük bir informasyon elde edilemez. Şu halde, eğer ancak üç soru sorarsak ($k=3$), elde edilen informasyon log 9 dan küçük olacaktır (en fazla 3 log 2=log 8 e eşittir); şu halde X in, hangi şehirde bulunduğunu ve hangi şehirli ile konuştuğunu öğrenmek için üç soru kâfi gelmiyecektir. Buna karşılık, akıllıca sorulan dört soru bunu temin etmeye kâfi gelecektir [gerçekten, $H(A_1) \leq 4 \log 2 = \log 16$ dır]. Şu dört sorunun, X in hangi şehirde bulunduğunu ve hangi şehirli ile konuştuğunu öğrenmesini mümkün kılacağı kolayca görülür:

1) A ve B şehirlerinin birisinde mi bulunuyorum?
2) C şehrinde mi bulunuyorum?
3) C şehrinde mi oturuyorsunuz?
4) A şehrinde mi bulunuyorum?

Gerçekten, 1) ve 2) sorularına verilen müspet veya menfi cevaplar soru sorulan kimsenin C şehrinden olduğunu gösterir; 3) sorusuna verilen bir menfi cevap, 2) sorusunun cevabının doğru olduğunu gösterir ve dördüncü soruya lüzum kalmaz; 3) sorusuna verilen bir müspet cevap, 1) sorusunun cevabının doğru olduğunu gösterir ve X in hangi şehirde bulunduğunu tayin etmek için 4) sorusunu sormak icap eder. 1) ve 2) sorularına verilen cevapların biri müspet, diğeri menfi olması hali soru sorulan kimsenin A yahut B şehirlerinin birisinden olduğunu gösterir, 3) sorusuna verilen menfi (yani (doğru) bir cevap soru sorulan kimsenin A şehrinde oturduğunu gösterir ve ancak 2) sorusuna verilen cevap menfi olduğu takdirde 4) sorusuna lüzum vardır; 3) sorusunun cevabı menfi (yani yanlış) olması, soru sorulan kimsenin B de oturdu-

ğunu gösterir ve ancak 2) sorusunun cevabı müspet ise 4) sorusu lüzumlu olur.

Şimdi aynı tipte diğer bir problemi inceliyelim.

PROBLEM 19. — Bir kimse 10 'dan (yahut 100 den yahut 1000 den veyahut herhangi bir N sayısından) küçük veya buna eşit bir pozitif tam sayı «düşünüyor». Bu sayıyı bulmak için kaç soru sormak lâzımdır (cevaplar «evet» veya «hayır» şeklinde olacaktır).

Neticesinin bulunması bahis konusu olan β deneyinin on mümkün neticesi vardır (düşünülen sayı 1 ile 10 arasında herhangi bir tam sayı olabilir). Soru sorulmadan önce tekmil neticeler eşit olarak muhtemeldir; β deneyinin $H(\beta)$ entropisi (başka ifade ile lüzumlu informasyon miktarı) $\log 10 = 1$ e eşittir. k sayıda soru sormaktan ibaret olan $A_k = \alpha_1\alpha_2, \cdots, \alpha_k$ karma deneyini gözönüne alalım. Bir soru sormaktan ibaret olan α_1 deneyinin entropisi $\log 2$ den küçük veya buna eşittir, zira α_1 in ancak iki mümkün neticesi vardır (biri müspet, diğeri menfî iki cevap). Şu halde A_k deneyinin entropisi $k \log 2$ den küçük veya buna eşittir (bak. evvelki problem). Bundan başka, β hakkında A_k deneyinde bulunan $I(A_k\ \beta)$ informasyonu, A_k da bulunan informasyondan küçük veya buna eşittir, başka ifade ile entropisi $H(A_k)$ dır. A_k deneyinin β nın neticesini tamamen belirlemesi için $I(A_k, \beta) = H(\beta)$ olması lâzımdır. Buradan

$$\log 10 = H(\beta) = I(A_k, \beta) \leq H(A_k) \leq k \log 2 = \\ = \log 2^k,\ 2^k \geq 10$$

yani

$$k \geq \log_2 10 = \frac{\log 10}{\log 2} \simeq 3{,}32$$

ve, k tam sayı olduğundan

$$k \geq 4$$

elde edilir.

Şimdi dört soru vasıtasiyle β nın neticesinin, yani x sayısının nasıl bulunacağını gösterelim. Evvelâ, ilk deneyde bulunan informasyonun mümkün olduğu kadar büyük olmasını, başka ifade ile entropisinin $\log 2$ ye eşit olmasını temin etmelidir. Bunun için, α_1 in iki neticesinin eşit olarak muhtemel olması lâzımdır. Bundan sonra, β hakkında α_1 de bulunan $I(\alpha_1, \beta)$ informasyonunun $H(\alpha_1)$ entropisine eşit olması şartı koşulur. Bunun için, birinci sorunun «yabancı» informasyon ihtiva etmemesi, yani $H_\beta(\alpha_1)$ bağlı entropinin sıfıra eşit olması veyahut β nın neticesinin α_1 in neticesini tamamen belirlemesi lâzımdır. Bu düşünceler ilk sorunun nasıl sorulacağını açıkça gösterir. x in tekmil değerlerini (yani 1 ile 10 arasında bulunan tekmil tam sayıları) iki eşit kısma ayıralım (α_1 in neticeleri eşit olarak muhtemel olmalıdır), ve x in bu iki kısımdan hangisinde bulunduğunu soralım; meselâ x in 5 den büyük olup olmadığı sorulabilir. Bu halde

$$I(\alpha_1, \beta) = H(\beta) - H_{\alpha_1}(\beta) = \log 2,$$

$$H_{\alpha_1}(\beta) = p(A_1) H_{A_1}(\beta) + p(A_2) H_{A_2}(\beta) = H(\beta) - \log 2$$

olacağı aşikârdır; burada A_1 ve A_2 ile α_1 in neticeleri gösterilmiştir: $p(A_1) = p(A_2) = 1/2$; yukarıdaki eşitliklerden

$$H_{A_1}(\beta) = H_{A_2}(\beta) = H(\beta) - \log 2$$

yazılır; böylece, α_1 in neticesi ne olursa olsun, β nın entropisi $\log 2$ kadar azalacaktır. Bundan sonra, kalan mümkün

halleri mümkün olduğu kadar aynı miktarda ihtiva eden iki kısma ayırmak ve x in hangi kısımda olduğunu belirlemek lâzımdır (x in 5 den büyük olduğu öğrenilmiş ise, x in 7 den büyük olup olmadığı sorulabilir; eğer x, 5 den küçük ise, x in 3 den büyük olup olmadığı sorulabilir) v.s. Eğer her defasında kalan mümkün haller, mümkün olduğu kadar eşit iki kısma ayrılırsa x, dört soruda tayin edilecektir *.

Keza, 100 yahut 1000 değer alabilen bir sayıyı bulmak için sorulacak en az soru sayısının sıra ile

$$k \geq \frac{\log 100}{\log 2} \cong \frac{2}{0{,}301} \quad \text{ve} \quad k \geqq \frac{\log 1000}{\log 2} \cong \frac{3}{0{,}301}$$

olduğu gösterilir; k tam sayı olduğundan

$$k \geq 7 \quad \text{ve} \quad k \geqq 10$$

elde edilir.

Genel olarak, *N sayıda değer alabilen bir x sayısını bulmak için sorulacak en az soru sayısı*

$$k - 1 < \frac{\log N}{\log 2} \leq k \quad (\text{yahut } 2^{k-1} < N \leq 2^k) \qquad (1)$$

---

*) *n tek sayı* (meselâ $n = 5$) olmak üzere, x in n mümkün değeri olduğunu belirledikten sonra α_{i+1} deneyi için eşit olarak muhtemel olan iki netice elde etmenin mümkün olmadığı aşikârdır: bu halde, mümkün değerler, aynı sayıda elemana malik olan iki kısma ayrılamaz; buradan anlazılır ki, α_{i+1} deneyinin $H(\alpha_{i+1})$ entropisi log 2 den küçük olacaktır. Bu, sorularımızın, informasyon bakımından, en uygunu olmadığını gösterir; başka ifade ile, aynı sayıda soru ile, biraz daha fazla miktarda mümkün değerler alabilen bir sayı bulmak mümkündür (meselâ, dört soru ile, $2^4 = 16$ değer alabilen bir sayıyı bulmak mümkün olabilir).

eşitsizlikleri ile belirlenmiştir. N sayısı ne olursa olsun

$$k \geq \frac{\log N}{\log 2}$$

olduğuna işaret edelim; üstelik, N sayısı yalnız 2 nin bir kuvveti olması (ve $\log N/\log 2$ nin tam sayı olması) halinde $k = \log N/\log 2$ dir. Bununla beraber N çok büyük ise k ile $\log N/\log 2$ arasındaki fark, bu sayılara nisbeten çok küçüktür (gerçekten, N büyük olduğu vakit $\log N/\log 2$ büyüktür, $k - \log N/\log 2$ farkı ise daima birimden küçük kalır). Şu halde *N nin çok büyük değerleri için,* gözönüne alınan β deneyinin $\log N$ entropisinin, sorulan bir sorudan ibaret olan α deneyinin ihtiva ettiği $\log 2$ informasyonuna oranı, β nın neticesini tayin etmek için sorulacak soru sayısı olan k yı verir.

Problem 19 ile aşağıdaki problemi karşılaştırmak faydalı olur:

PROBLEM 20. — Bir kimse, N sayıda değer alabilen bir x sayısı «düşünüyor»; bu kimsenin, sorulan sorulara «evet» yahut «hayır» ile cevap vermesi şartı ile, bu x sayısını sorular sorarak belirlenmesi bahis konusudur; üstelik soru soranın x sayısı hakkında bir informasyona sahip olduğu ve bu suretle N sayıda değerin eşit olarak muhtemel olduğu kabul edilemiyeceği farz ediliyor; başka ifade ile bu değerlerden birinin aranılan sayı olması şansı diğerlerinden daha fazladır *. Bu halde sorular nasıl sorulmalıdır?

*) Meselâ, düşünülen sayının bir kâğıda yazıldığı ve soru soracak kimsenin bunu gördüğü fakat hangi sayı olduğundan tamamen emin olmadığı düşünülebilir.

Mâlik olduğumuz informasyondan dolayı mümkün hallerden hiç biri tamamen hariç bırakılmıyor ise (aksi halde mümkün hallerin N sayısı yerine daha küçük bir sayı almak icap edecektir), tekmil halerde x sayısını bulmayı mümkün kılan en az soru sayısının daima (1) eşitsizlikleri ile verildiği aşikârdır; bundan başka sorular yukarıda gösterildiği gibi sorulmalıdır. Gerçekten, tekmil hallerde soruların cevaplarından müstakil olarak daha az sayıda soru kâfi gelseydi bu problem 19 un neticesi ile tezat teşkil ederdi. Fakat bu, daima, tekmil neticeler eşit olarak muhtemel değilmiş gibi hareket etmek lâzımdır demek değildir. Meselâ, aranılan sayının bir x_0 değerine mâlik olması şansı çok büyük ise (meselâ bu neticenin ihtimali 0,99 a eşit veya daha fazla ise), menfî bir cevap halinde fazla bir informasyon elde etmeden bir soru harcanmasına rağmen (mümkün değerler sayısı ancak bir eksilir), tabiatiyle evvelâ x in x_0 a eşit olup olmadığını sormak lâzımdır. Genel olarak, *her defasında x in mümkün değerlerini iki kısma ayırmak lâzımdır, şöyle ki, x in bu kısımlardan birine veya diğerine ait olması ihtimalleri mümkün olduğu kadar birbirine yakın olsun.* Bu, x in iki kısımdan hangisine ait olduğunu belirlemekten ibaret olan α deneyine bir maksimum entropi ve β deneyi hakkında α da bir maksimum informasyon temin eder. Bu tarzda, en uygun olmayan halde *lüzumlu soruların minimum sayısının* ne olacağını bilemeyiz, fakat sorulan soruların *ortalama değeri* herhangi tarzda soru sorma halindekinden küçük olacaktır.

Bu son iddiayı ispat etmeden sadece bir özel halde gerçekleyeceğiz *. x in mümkün değerlerinin N sayısı 4 e

*) Problem, N nin büyük olması ve neticelerin herbirinin ihtimalleri çok küçük olması halinde bütün açıklığı ile incelenebilir; bu halde lüzumlu soru sayısı $H(\beta)/\log 2$ ye eşittir (bak. § 2, bölüm IV).

eşit olsun; bu halde, (1) eşitsizlikleri ile belirlenen k sayısı 2 ye eşittir. x değerinin diğer x_1, x_2, x_3 değerlerinden daha muhtemel olduğunu farz edelim; x in x_0 a eşit olması ihtimali p, ve x_i ye eşit olması ihtimali q olsun (i, burada 1, 2, 3 değerlerinden birini alabilir; $p > q$, $p + 3q = 1$ dir). Birinci soru olarak x in x_0 yahut x_1 değerlerinden birine eşit olup olmadığı sorulabilir. Derhal x in, x_0 eşit olup olmadığı da sorulabilir. Bu iki deneyi $\alpha_1^{(1)}$ ve $\alpha_1^{(2)}$ ile gösterelim; $\alpha_1^{(1)}$ deneyinin iki neticesinin ihtimallerinin $p + q$ ve $2q$ olduğu aşikârdır [ve böylece $H(\alpha_1^{(1)}) = -(p+q)\log(p+q) - 2q\log 2q$ dür]; $\alpha_1^{(2)}$ nin iki neticesinin ihtimalleri ise p ve $3q$ dür [$H(\alpha_1^{(2)}) = -p\log p - 3q\log(3q)$]. Eğer $p > 1/2$ ise, $\alpha_1^{(2)}$ in neticelerinin ihtimalleri $\alpha_1^{(1)}$ in neticelerinkine nazaran daha birbirine yakındır; $1/2 > p > q$ ise, $\alpha_1^{(1)}$ ve $\alpha_1^{(2)}$ nin neticelerinin ihtimallerinin $(p+q) - 2q = p - q$ ve $3p - q$ farklarını karşılaştırmak lâzımdır. $p > 2q$ yani yani $p > 2/5$ ise $3q - p < p - q$ olduğundan [gerçekten $q > 2/5$ ise $q = 1/3(1-p)$ ve $p > 2/3(1-p)$ dir], $p > 2/5$ ise $\alpha_1^{(2)}$ deneyi ile ve $p < 2/5$ ise $\alpha_1^{(1)}$ deneyi ile başlamak lâzımdır.

Eğer «x, x_0 yahut x_1 sayılarından birine eşit mi?» sorusu ile başlarsak, *x in mümkün değerlerini, ihtiva ettikleri değer sayıları eşit olan* iki kısma ayırırız; bu halde ilk sorunun cevabı ne olursa olsun x i iki soruda bulacağız. Eğer «x, x_0 a eşit mi?» sorusu ile başlarsak, x i bir soru ile belirleme şansımız vardır, bu ihtimal x in x_0 a eşit olması ihtimali yani p dir. Bununla beraber, x, x_0 a eşit değilse x i ikinci soru il bulabileceğimizden emin olamayız; «x, x_1 e eşit mi?» sorusu (q ihtimali ile) müspet olabilir yahut (x in x_2 yahut x_3 e eşit olması ihtimali ile yani $2q$ ye eşit bir ihtimal ile) menfî olabilir; bu sonuncu halde bir üçüncü soru daha lâzımdır. Şu halde, $\alpha_1^{(2)}$ deneyi

ile başlanması halinde, x i bir soru ile belirleme ihtimali p, iki soru ile belirleme ihtimali q, üç soru ile belirleme ihtimali $2q$ dür. Buradan anlaşılır ki, sorulacak soru sayısının ortalama değeri

$$p \cdot 1 + q \cdot 2 + 2q \cdot 3 = p + 8q$$

ye eşittir. Eğer $p > 2/5 (p + 8q) = 1/2 (8 - 5p)$ ise $p + 8q < 2$ olduğunu gerçeklemek kolaydır, zira $1/3q = (1 - p)$ dir; şu halde $p > 2/5$ olması halinde $\alpha_1^{(1)}$ deneyi ile başlamak yersiz değildir.

§ 2. Kalp madenî paraları tartı yolu ile belirleme Problemleri

Evvelâ, problem 19 a çok benzeyen aşağıdaki problemi inceleyelim.

PROBLEM 21. — 25 tane aynı değerde madenî para olsun; bunların 24 tanesi aynı ağırlıktadır; biri kalp olup diğerlerine nazaran biraz hafiftir. Bir terazi yardımı ile ve ağırlık kullanmadan, en az kaç tartı ile kalp paranın bulunabileceği soruluyor.

Neticesinin belirlenmesi bahis konusu olan β deneyinin 25 mümkün neticesi vardır (kalp para 25 taneden biri olabilir); tekmil bu neticeler tabiatiyle eşit olarak muhtemel olduğundan, $H(\beta) = \log 25$ dir. Başka ifade ile, kalp parayı tayin etmek için $\log 25$ e eşit informasyona lüzum vardır. (Herhangi) bir tartı yapmaktan ibaret olan α_1 deneyinin üç mümkün neticesi vardır (terazi dengede olmaz, ya sağdaki kefe ağır gelir yahut soldaki kefe ağır gelir, veyahut terazi kefeleri dengededir); şu halde bu

deneyle elde edilen informasyon $H(\alpha_1) \leq \log 3$ dür. k sayıda tartma işlminden ibaret olan karma deney $A_k = \alpha_1 \cdot \alpha_2 \cdots \alpha_k$ olsun; bu deney, $k \log 3$ den küçük veya buna eşit bir informasyon verir (bak. sahife 100). A_k deneyinin β nın neticesini tamamen belirlemesi için

$$H(A_k) \geqq H(\beta) \qquad \text{yahut} \qquad k \log 3 \geq \log 25$$

olmalıdır. Buradan $3^k \geq 25$ yani

$$k \geqq \log_3 25 = \frac{\log 25}{\log 3}$$

yahut, k tam sayı olduğundan

$$k \geq 3$$

neticesi bulunur.

Kalp paranın üç tartma işlemi ile hakikaten bulunabileceğini göstermek kolaydır. α_1 deneyinin vukuu ile elde edilen informasyonunun mümkün olduğu kadar büyük olması için, bunun üç neticesinin ihtimallerinin mümkün olduğu kadar birbirlerine yakın olması lâzımdır. Terazinin kefelerinden her birine m tane madenî para konduğunu farz edelim (kefelere farklı sayılarda para koyma faydasızdır, zira bu halde tartma işleminin neticesi önceden belli olup elde edilen informasyon sıfıra eşittir); şu halde kefelerin dışında $25 - 2m$ madenî para kalır. Kalp paranın n tane para içinde bulunması ihtimali $n/25$ olduğundan, α_1 in üç neticesinden herbirinin ihtimalleri, sıra ile, $m/25$, $m/25$ ve $1 - (2m/25)$ dir, bu ihtimaller $m = 8$ ve $25 - 2m = 9$ olması halinde en fazla birbirlerine yakın olurlar. Şu halde kefelerden her birine 8 er para koyarak ilk tartma işlemi (α_1 deneyi) kalp parayı

ihtiva eden 9 tanelik bir para gurupu (denge olması halinde), yahut 8 tanelik bir para gurupu (kefelerden biri ağır gelmesi halinde) bulmayı mümkün kılar. Her iki halde de, kefelerden her birine bu gurupun paralarından 3 er tane koymak suretiyle yapılan ikinci tartma işlemi (α_2 deneyi) bize en büyük informasyonu verir. Şu halde $\alpha_1\alpha_2$ karma işlemi aralarında kalp parayı ihtiva eden 3 (yahut 2) tanelik bir para gurupunu bulmayı mümkün kılar. Üçüncü tartma işleminde (α_3 deneyi) kefelerden her birine bu son gurupun paralarından birer tane koymak suretiyle kalp parayı kolayca tayin ederiz.

Keza, *N sayıda bir para gurupu içinde bulunan bir kalp parayı (diğerlerinden hafif olanı) bulmak için en uygun olmayan halde lüzumlu minimum tartma işlemi sayısı olan k nın*

$$3^{k-1} < N < 3^k \quad \text{yahut} \quad k-1 < \frac{\log N}{\log 3} < k \qquad (2)$$

eşitsizlikleri ile verildiği gösterilir.

Eğer N çok büyük ise bu k sayısı çok iyi bir yaklaşımla $\log N/\log 3$ oranı ile, başka ifade ile, kalp parayı belirtmekten ibaret olan β deneyinin entropisinin, bir tartma işlemi yaparak elde edilebilen maksimum entropiye oranı ile verilmiştir (bak. sahife 103).

İleride, biraz daha genel bir probleme ait benzer bir netice ile ilgileneceğiz. Evvelâ şuna işaret edelim ki, biri kalp olan ve diğerlerinde biraz *ağır* olan N tane madenî para varsa, bu kalp parayı belirlemek için ağırlık kullanmadan bir terazi ile yapılan tartma işlemlerinin minimum sayısı yine (2) eşitsizlikleri ile verilmiştir: hafif bir para yerine ağır bir para alınması ispatlarımızda hiç bir deği-

şiklik yapmaz. Şimdi daha genel bir hali inceleyelim; *N tane madeni para iki gurupa ayrılmış olsun: bir A gurupunda a sayıda para ve bir B gurupunda* $b = N - a$ *sayıda para bulunsun; N tane paradan birisinin kalp olduğunu ve bu kalp para A gurupunda ise diğerlerinden hafif olduğunu biliyoruz; kalp parayı belirlemek için yapılması gereken minimum tartma işlem sayısının yine (2) eşitlikleri ile verildiğini* gösterelim *; $b = 0$ hali, evvelki problemin sonundaki haldir.

Bizi ilgilendiren β deneyinin N sayıda mümkün neticesi olduğundan $3^k \geq N$ dir, aksi halde, k sayıda tartma işleminden ibaret olan $A_k = \alpha_1 \cdot \alpha_2 \cdots \alpha_k$ deneyi β deneyinin neticesini belirleyemez (filhakika, bu halde

$$I(A_k, \beta) \leq H(A_k) \leq k \log 3 = \log 3^k < \log N = H(\beta)$$

dır; β nın neticeleri daima eşit olarak muhtemel farz edilmiştir). Diğer taraftan $N \leq 3^k$ ise kalp parayı daima k sayıda tartma işlemi ile bulmak mümkündür; bunu rekürans yolu ile göstermek kolaydır. Gerçekten, $k = 1$ için yani $N = 1$, 2 yahut 3 için iddiamız (evvelki notta gösterilen istisnaî hal dışında) aşikârdır; $N = 1$ için kalp para önceden bellidir; $N = 2$ ($a = 2$ yahut $b = 2$) ve $N = 3$ için, gruplardan birinin iki parasını karşılaştırmakla kalp para belirlenir. $N \leq 3^k$ için, k tartma işlemi ile kalp paranın bulunmasının mümkün olduğunu kabul edelim ve $3^k < N \leq 3^{k+1}$ olsun. A gurupundaki paralardan $2x$ çift sayıdakini ve B gurupundaki paralardan $2y$ çift sayıdakini ayıralım, şöyle ki,

$$x + y \leq 3^k, \quad N - (2x + 2y) \leq 3^k$$

*) $N = 2$ istisnaî halinde $a = b = 1$ dir ve kalp parayı bulmak mümkün olmaz.

yani

$$3^k \geqq x + y \geq \frac{N - 3^k}{2}$$

olsun. Şimdi terazinin kefelerinden her birine x sayıda A gurupundaki paralardan ve y sayıda B gurupundaki paralardan koyalım; geriye $N_1 = N - 2x - 2y \leq 3^k$ tane para kalacaktır. Bu tartma işleminde (α_1 deneyi) terazi kefeleri dengede kalırsa, kalp para geri kalan N_1 tane para arasında (yani A gurupunun $a_1 = a - 2x$ paraları yahut B gurupunun $b_1 = b - 2y$ paraları arasında) bulunur; kefelerden biri diğerinden ağır gelirse, kalp para ya daha hafif gelen kefedeki A gurupunun x paraları arasında yahut daha ağır gelen kefedeki B gurupunun y paraları arasında bulunur. Fakat $N_1 \leq 3^k$ ve $x + y \leq 3^k$ olduğundan, hipotezimize göre, k sayıda tartma işlemi ile her halde, kalp parayı bulabiliriz *; buradan, $N < 3^{k+1}$ sayıda paralar arasından kalp paranın en fazla $k + 1$ tartma işlemi ile bulunabileceği anlaşılır. Bu ise yukardaki iddiayı gösterir.

Aynı problem, N sayıda tekmil paraların kalp olma ihtimalinin aynı olmaması halinde sorulabilir (meselâ bunlardan bazılarının dış görünüşü bizi şüpheye düşürebilir). Promlem 20 de (sahife 104) olduğu gibi, paraları, kalp paranın terazinin iki kefesine konulan eşit sayıda para ihtiva eden iki gurupla bir kenara bırakılan gurupta bulunması ihtimalleri mümkün olduğu kadar birbirine

*) $N > 2$ ise $x = y = 1$ yahut $a_1 = b_1 = 1$ hali artık istisnai hal değildir; gerçekten, A grupunun şüpheli parasından ve B grupunun şüpheli parasından başka, elimizde, kalp olmadığından emin olduğumuz belirli sayıda para bulunur; bu paralardan birinin ağırlığı, şüpheli paralardan birinin ağırlığı ile kaşılaştırılarak, yalnız bir tartma ile, kalp para belirlenebilir.

yakın olan, üç gurupa ayırmak abes bir işlem olmaz. Böyle bir işlem yapılırsa kalp parayı bulmak için lüzumlu tartma işlemi sayısı belki, (2) eşitsizlikleri ile verilen sayıdan büyük olabilir; bununla beraber, bu tartma işlemi sayısının *ortalama değeri* mümkün olduğu kadar küçük olacaktır.

Şimdi aynı tipte fakat evvelkine nazaran biraz daha karışık bir problemi inceleyeceğiz.

PROBLEM 22. — 12 tane madeni para gözönüne alalım; bunların 11 i aynı ağırlıkta olsun; on ikincisi kalp olup diğerlerinden ancak ağırlığı ile farklıdır (üstelik diğerlerinden daha hafif veya daha ağır olduğu da bilinmiyor). Kalp parayı bulmak için bir terazi ile yapılan lüzumlu tartma işleminin minimum sayısı nedir? Kalp para diğerlerinden daha hafif mi yoksa ağır mıdır?

Aynı problem para sayısı 13 olması halinde de inceleyiniz.

Burada β deneyinin 24 (yahut 26) mümkün neticesi vardır (12 yahut 13 madeni paranın her biri kalp olabilir, ve üstelik, herbiri diğerlerinden daha hafif veya daha ağır olabilir). Tekmil neticelerin eşit olarak muhtemel oldukları kabul edilirse, β nın $H(\beta)$ entropisi $\log 24$ e yahut $\log 26$ ya eşit olacaktır. Şu halde $\log 24$ (yahut 26) a eşit bir informasyon elde edilmelidir. k sayıda tartma işleminden ibaret olan $A_k = \alpha_1\alpha_2 \cdots \alpha_k$ karma işlemi $k \log 3 = \log 3^k$ ya eşit olan bir maksimum informasyon verdiğinden ve $3^3 = 27$ olduğundan, ilk bakışta kalp parayı bulmak ve bunun diğerlerinden hafif veya ağır olduğunu belirlemek için, para sayısının 12 veya 13 olması hallerinde, 3 tartma işleminin kâfi geleceği düşünülür; hakikatte 3 tartma işlemi, para sayısının 13 olması halinde

kâfi değildir; bu keyfiyet, 3 tartma işleminin verdiği informasyon biraz daha dikkatle incelenerek kolayca görülür.

Gerçekten, ilk tartmada terazi kefelerinden her birine 1 er, yahut 2 şer, yahut 3 er, yahut 4 er, yahut 5 er ve nihayet yahut 6 şar para konabilir. Bu deneyleri $\alpha_1^{(i)}$ ile gösterelim $(i=1, 2, \cdots, 5, 6)$. Eğer i sayısı 1, 2, 3 yahut 4 e eşit ise ve eğer terazi dengede ise, $\alpha_1^{(i)}$ deneyi, kalp paranın, geri kalan $13-2i$ sayıda paralardan biri olduğunu gösterir; bu sayı en az 5 e eşit olduğundan, geriye en az 10 mümkün netice kalır ve sonra yapılacak iki tarma işlemi kalp parayı ve ağırlığını belirlemeye kâfi gelmeyebilir (gerçekten, $2 \log 3 = \log 9 < \log 10$ dur). Eğer i sayısı 5 yahut 6 ise ve eğer $\alpha_1^{(i)}$ deneyinde terazinin kefelerinden biri (meselâ sağdaki) diğerlerinden daha ağır gelirse $\alpha_1^{(i)}$ deneyi, kalp para diğerlerinden daha ağır ise, sağdaki kefe içindeki i sayıda paralardan biri olduğunu, eğer kalp para diğerlerinden daha hafif ise, soldaki kefedeki i sayı da paralardan biri olduğunu gösterir; bu halde de, β deneyi için $i+i=2i \geq 10$ mümkün netice kalır ve iki tartı işlemi neticeyi belirlemek için yine kâfi gelmez.

Şimdi para sayısının 12 olması halini inceliyelim. $\alpha_1^{(i)}$ ilk deney olsun, burada terazinin her kefesine i sayıda para koyuyoruz. Eğer terazi dengede kalırsa ($\alpha_1^{(i)}$ nin bu neticesini E ile göstereceğiz), kalp para geri kalan $12-2i$ sayıda paralar arasında bulunur, bu ise β nın (24 toplam neticesinden) $2(12-2i)$ mümkün neticesine tekabül eder. Eğer sağdaki kefe daha ağır gelirse (D neticesi), kalp para diğerlerinden ağır ise sağ kefedeki i sayıda paralar arasında, diğerlerinden hafif ise sol kefedeki i sayıda paralar arasında bulunur; şu halde bu,

β nın $2i$ sayıda neticesine tekabül eder. Keza, eğer soldaki kefe daha ağır gelir ise (G neticesi), bu, β nın $2i$ sayıda neticesine tekabül eder. Demek oluyor ki, $\alpha_1^{(i)}$ nin üç neticesinin ihtimalleri sıra ile

$$\frac{2(12-2i)}{24}=\frac{6-i}{6},\ \frac{2i}{24}=\frac{i}{12}\ \text{ve}\ \frac{i}{12}$$

dir. Buradan anlaşılır ki $\alpha_1^{(1)}$, $\alpha_1^{(2)}$, $\alpha_1^{(3)}$, $\alpha_1^{(4)}$, $\alpha_1^{(5)}$ ve $\alpha_1^{(6)}$ deneylerinden entropisi en büyük olan $\alpha_1^{(4)}$ deneyidir, zira bu deneyin üç neticesi eşit olarak muhtemeldir; şu halde bize en büyük informasyonu veren bu deneydir ve bu deney ile başlamak en uygundur. Şimdi iki mümkün hali inceliyelim.

A) Bu ilk tartmada, kefeler dengede kalır. Şu halde kalp para geri kalan dört para arasındadır. O halde iki tartma ile bunlardan hangisinin kalp olduğunu ve kalp paranın diğerlerinden hafif veya ağır olduğunu belirlemek lâzımdır; $2\cdot 4=8$ mümkün hal kaldığından, ve $3^2=9>8$ olduğundan bunun mümkün olacağı ümit edilebilir. Bununla beraber terazinin kefelerinin her birine bu dört paradan, ikisini bir kenara bırakarak, birer tane koyarsak ($\alpha_2^{(1)}$ deneyi), ve eğer terazi dengede kalırsa, hâlâ mümkün olan dört neticeden hangisinin vuku bulduğunu bir tartma işlemi ile belirlemek lâzımdır ve bu ise mümkün değildir (zira $4>3$ dür); eğer dört parayı ikişer ikişer kefelere koyarsak ($\alpha_2^{(2)}$ deneyi) ve kefelerden biri daha ağır gelirse, yine β deneyi için dört mümkün hal kalır ve yine kalp parayı belirlemek için iki tartma işlemi kâfi gelmez. Şu halde para sayısı 12 olması halinde problemi çözmek için üç tartma işlemi kâfi gelmez gibi görünür.

Fakat acele hüküm vermemek lâzımdır. Zira henüz elimizde ikinci tartıda kullanmadığımız ve kalp olmayan $4+4=8$ tane para bulunmaktadır; şu halde α_2 deneyi için ikiden fazla mümkün hal vardır. Elimizdeki dört tane paradan sağdaki kefeye i tane soldakine $j \leq i$ tane ve bundan başka soldaki kefeye kalp olmadığı bilinen paralardan $i-j$ tane para koymaktan ibaret olan deneyi $\alpha_2^{(i,j)}$ ile gösterelim; bu tarzda, $\alpha_2^{(1,1)}$ ve $\alpha_2^{(2,2)}$ deneyleri yukarıda incelenen $\alpha_2^{(1)}$ ve $\alpha_2^{(2)}$ deneyleridir. $p(E)$, $p(D)$, $p(G)$ ile, $\alpha_2^{(i,j)}$ deneyinde sıra ile, terazinin dengede kalması, sağ kefenin ağır gelmesi, sol kefenin ağır gelmesi ihtimallerini gösterelim. Bu ihtimalleri hesaplamak kolaydır; bunlar sıra ile E, D ve G olan uygun neticelerin, toplam neticeler (8 e eşit olan) sayısına oranına eşittir. Mütekabil değerler aşağıdaki cetvelde gösterilmiştir, bu cetvel aynı zamanda $\alpha_2^{(i,j)}$ deneyinin $H(\alpha_2^{(i,j)})$ entropisini de verir [bu entropi $-p(E)\log p(E) - p(D)\log p(D) - p(G)\log p(G)$ ye eşittir]:

i	j	$p(E)$	$p(D)$	$p(G)$	$H(\alpha_2^{(i,j)})$
1	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0,452
1	0	$\frac{3}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	0,320
2	2	0	$\frac{1}{2}$	$\frac{1}{2}$	0,301
2	1	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{3}{8}$	0,470
2	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0,452
3	1	0	$\frac{1}{2}$	$\frac{1}{2}$	0,301
3	0	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{3}{8}$	0,470
4	0	0	$\frac{1}{2}$	$\frac{1}{2}$	0.301

Bu cetvelde, $\alpha_2^{(2,1)}$ ve $\alpha_2^{(3,0)}$ deneylerinin en büyük entropiye malik oldukları görülür ve bu entropi 0.470 e eşittir; şu halde maksimum informasyonu elde etmek için kefelerden birine, içinde kalp para bulunması muhtemel olan dört şüpheli paradan ikisi, ve diğer kefeye bir şüpheli para ve bir de kalp olmadığı bilinen bir para, yahut kefelerden birine üç şüpheli para ve diğer kefeye üç tane kalp olmayan para koymak lâzımdır. Bu iki halde de üçüncü tartma ile kalp paranın ve ağırlığının belirlenebileceğini görmek kolaydır. Gerçekten $\alpha_2^{(2,1)}$ yahut $\alpha_2^{(3,0)}$ deneyleri E neticesini verirlerse kalp para, tartılmaya katılmayan şüpheli paradır; diğerlerinden ağır yahut hafif olduğunu tayin etmek için 11 paranın biri ile karşılaştırmak kâfidir (üçüncü tartı). Eğer $\alpha_2^{(2,1)}$ deneyinin neticesi D ise kalp para, ya terazinin sağ kefesinde bulunan iki paradan biridir ve bu halde diğerlerinden ağırdır, yahut geri kalan tek şüpheli paradır ve bu halde diğerlerinden hafiftir; sağ kefede bulunan iki paranın ağırlıkları karşılaştırılarak (üçüncü tartma) problemin çözümü bulunacaktır (iki para aynı ağırlıkta ise kalp para, üçüncü şüpheli paradır; aksi halde ikisinden ağır olanıdır). Eğer $\alpha_2^{(3,0)}$ deneyinin neticesi D ise kalp para sağ kefedeki üç şüpheli paradan biridir, üstelik diğerlerinden ağırdır; bu üç şüpheli paradan ikisinin ağırlıkları karşılaştırılarak (üçüncü tartı) β nın neticesi bulunur (ikisi de aynı ağırlıkta ise kalp para üçüncü şüpheli paradır, aksi halde, iki şüpheli paradan ağır olanıdır). Keza $\alpha_2^{(2,1)}$ yahut $\alpha_2^{(3,0)}$ deneylerinin neticesi G olması hali incelenebilir.

B) Birinci deneyde, kefelerden birinin (meselâ sağdakinin) daha ağır geldiğini farz edelim. Bu halde kalp para ya sağ kefedeki paralardan biridir ve bu takdirde diğerlerinden ağırdır, yahut soldaki dört paradan biridir

ve bu halde ise diğerlerinden hafiftir. İkinci tartmada, sağdaki kefeye, «sağın» paralarından i_1 tane «solun» paralarından i_2 tane, ve soldaki kefeye, «sağın» paralarından j_1 tane «solun» paralarından j_2 tane koyabiliriz ve, kalp olmadığını bildiğimiz $(i_1 + i_2) - (j_1 + j_2)$ tane para ile (bunlar ilk tartıya katılmayan dört paradır) kefelerdeki paraların sayısını eşit kılarız; $i_1 + i_2 \geqq j_1 + j_2$ farz ediyoruz. Buna tekabül eden deneyi $\alpha_2^{(i_1, i_2; j_1, j_2)}$ ile göstereceğiz. Burada da yukarıda yapıldığı gibi i_1, i_2, j_1 ve j_2 nin tekmil mümkün değerleri ile $\alpha_2^{(i_1, i_2; j_1, j_2)}$ deneylerinin entropilerini belirlemek mümkündür; fakat mümkün hallerin sayısı kâfi derecede büyük olduğundan bunlardan bazılarını önceden gözönüne almama yolunu aramak daha uygun olur. Üçüncü tartının (α_3 deneyinin) yapılması ile elde edilen informasyon log 3 den büyük olamıyacağından [gerçekten $H(\alpha_3) \leq \log 3$ dür], ikinci tartıdan sonra β için üç mümkün neticeden fazlasının olmaması icap ettiğine işaret edelim. Buradan, ikinci tartıya katılmayan şüpheli para sayısının 3 den büyük olamıyacağı neticesi çıkarılır, zira α_2 nin neticesi E olursa bu paralar yine şüpheli kalırlar.

Şu halde

$$8 - (i_1 + i_2 + j_1 + j_2) \leq 3, \quad i_1 + i_2 + j_1 + j_2 \geq 5$$

olmaması lâzımdır ve

$$i_1 + i_2 \geq j_1 + j_2$$

olduğundan

$$i_1 + i_2 \geq 3$$

olmalıdır.

Eğer $\alpha_2^{(i_1, i_2; j_1, j_2)}$ deneyinin neticesi D ise kalp para ya sağ kefedeki «sağın» i_1 paralarından biridir ve bu halde diğerlerinden ağırdır, yahut sağ kefedeki «solun» j_2 paralarından biridir ve bu halde diğerlerinden hafiftir; keza, bir G neticesi halinde, kalp para ya sağ kefedeki «solun» i_2 paralarından biridir, yahut sol kefedeki «sağın» j_1 paralarından biridir. Şu halde iki eşitsizlik daha elde ederiz:

$$i_1 + j_2 \leq 3,$$

$$i_2 + j_1 \leq 3.$$

Şimdi şartlarımızı sağlayan tekmil halleri aşağıdaki cetvelde gösterelim:

i^1	i_2	j^1	j_2	$p(E)$	$p(D)$	$p(G)$	$H(\alpha_2^{(i_1, i_2; j_1, j_2)})$
2	1	2	1	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{3}{8}$	0,470
2	1	2	0	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	0,470
2	1	1	1	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{4}$	0,470
1	2	1	2	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{3}{8}$	0,470
1	2	1	1	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	0,470
1	2	0	2	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{4}$	0,470
3	1	1	0	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{4}$	0,470
2	2	1	1	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{3}{8}$	0,470
2	2	1	0	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	0,470
2	2	0	1	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{4}$	0,470
1	3	0	1	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	0,470
3	2	1	0	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{3}{8}$	0,470
2	3	0	1	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{3}{8}$	0,470

Böylece, β deneyi hakkında, α_2 deneyinin 0,470 e eşit informasyon ihtiva eden, evvelkinde olduğu gibi 2 değişik hal değil, 12 değişik halin mevcut olduğunu görüyoruz [burada $I(\alpha_2, \beta)$ informasyonunun $H(\alpha_2)$ ye eşit olduğu tamamen aşikârdır]. Seçilen hal hangisi olursa olsun, elde edilen informasyon miktarı β nın neticesini bir deney ile (üçüncü tartma ile) belirlemeye kâfi gelecektir. Meselâ $\alpha_2^{(2,1;2,1)}$ deneyi için bir *E* neticesi olması halinde kalp para, tartmaya katılan «soldaki» iki paradan biridir; bunun diğerlerinden daha hafif olduğu bilindiğinden, bu iki paranın ağırlıklarını karşılaştırmak kâfidir (yahut bu iki paradan birini, kalp olmadığı bilinen bir para ile de karşılaştırılabilir. Aynı deneyin bir *D* neticesi halinde, kalp para diğerlerinden ağırdır ve sağ kefedeki «sağın» iki parasından biridir, yahut kalp para diğerlerinden hafiftir ve bu, sol kefede bulunan «solun» tek parasıdır; o halde kalp parayı bulmak için «sağın» iki şüpheli parasının ağırlıklarını karşılaştırmak kâfidir. $\alpha_2^{(2,1;2,1)}$ deneyinin *G* neticesini vermesi hali aynı tarzda incelenir.

Böylece problem, para sayısı 12 olması halinde tamamen çözülmüş olur. Şimdi para sayısının 13 olması halinde, kalp paranın belirlenmesi için dört tartma işleminin kâfi geleceğini göstermek kolaydır (yukarıda üç tartmanın yetersizliğini göstermiştik). Terazi kefelerinin her birine 4 er para koyalım ve 5 tanesini bir kenara bırakalım. Kefelerden biri ağır gelirse, para sayısının 12 olması problemindeki ilk tartmanın bir *D* neticesinin vermesi halindeki aynı durumda bulunuyoruz (ehemmiyeti olmayan yegâne fark, kalp olmadığından emin olduğumuz para sayısının 4 yerine burada 5 olmasıdır); bu halde, β nın neticesini üç tartma ile belirlemek mümkündür. Terazi dengede kalırsa, şüpheli para sayısı 4 yerine 5 dir. Bu halde şüpheli

paralardan birinin ağırlığını kalp olmadığı bilinen paralardan biri ile karşılaştırabiliriz; terazi dengede kalmazsa problem çözülmüştür; terazi dengede ise 4 şüpheli para ile karşı karşıya bulunuyoruz; bu takdirde, iki tartma ile hangi paranın kalp olduğu ve ağırlığı belirlenebilir (bak. A hali, sahife 114).

Problem 22 nin şartlarında, kalp paranın bulunması ve bunun *diğerlerinden daha ağır veya daha hafif* olduğunun belirlenmesi istenmektedir. Bu son şarttan sarfı nazar edilirse, birçok hallerde lüzumlu tartma sayısının eksileceği ilk bakışta ümit edilebilir. *Gerçekten* β nın mümkün neticelerinin sayısı yarıya iner (para sayısı N olduğuna göre mümkün neticeler $2N$ yerine N olur), ve deneyin entropisi hayli ($\log 2$ kadar) azalır. Para sayısı N nin $3^3 = 27$ olması halinde, yahut N nin 27 den biraz küçük olmasına tekabül eden her halde, kalp paranın üç tartma işlemi ile bulunacağı ümit edilebilir. Hakikatte bu böyle olmaz: para sayısının 13 olması halinde kalp para üç tartma ile belirlenebilirsee de bu, para sayısı 14 olması halinde mümkün değildir ve bu halde dört tartma işlemi lâzımdır. Bu, şundan ileri gelir: ağırlığı aranmadan, sadece kalp para aranıyorsa, bir terazide ağırlık kullanmadan paraları tartmaktan ibaret olan α deneyinin entropisi, genel olarak $I(\alpha, \beta)$ informasyonuna eşit olmayacaktır. Zira, artık β nın neticesi α nın neticesini belirlemez: hangi kefenin ağır geleceğini bilmek için, kalp paranın hangi kefede bulunduğunu bilmek kâfi gelmez, bu kalp paranın diğerlerinden daha ağır veya daha hafif olduğunu da bilmek lâzımdır. α deneyinin bir «yabancı» informasyon (bak. sahife 81), daha açık olarak, kalp paranın ağırlığı hakkında bir informasyon ihtiva ettiği de söylenebilir. Bundan dolayı bu halde, lüzumlu olan tartma işlemlerinin

sayısı $\log N/\log 3$ ün hayli üzerinde bulunur ($\log N$, β deneyinin entropisidir; $\log 3$ bir tartmanın maksimum entropisidir).

Evvelki problemin genelleştirilmiş hali:

PROBLEM 23. — Aynı değerde N tane madenî para gözönüne alalım; bunlardan biri kalp olup diğerlerinden ya biraz daha ağır yahut biraz daha hafiftir. Kalp parayı bulmak ve bunun diğerlerinden daha hafif veya daha ağır olduğunu belirlemek için bir terazide ağırlık kullanmadan yapılan tartma işlemlerinin minimum k sayısı nedir?

Evvelâ, gözönüne alınan deneyin entropisi $\log 2N$ ye eşit olduğundan ve k sayı da tartma işleminden ibaret olan $A_k = \alpha_1, \alpha_2, \cdots, \alpha_k$ deneyinin entropisi $k \log 3 = \log 3^k$ dan küçük olduğundan

$$2N \leq 3^k, \quad N \leq \frac{3^k}{2}$$

dir; şu halde N ve k tam ve 3^k tek sayı olduğundan

$$N \leq \frac{3^k - 1}{2}$$

olur. Demek oluyor ki,

$$k \geq \log_3 (2N + 1) = \frac{\log (2N + 1)}{\log 3}$$

dir. Meselâ $N > 1/2\ (3^3 - 1) = 13$ ise, kalp para üç tartma işleminden daha azı ile belirlenemez.

Hattâ, $N = 1/2\,(3^k - 1)$ ise, kalp parayı bulmak ve ağırlığını belirlemek için k sayıda tartmanın kâfi gelmediğini görmek kolaydır (meselâ, $N = 13$ için bu belirleme tekmil hallerde mümkün değildir). Bu keyfiyetin ispatı, $N = 13$ için verilmiş olan ispattan esas bakımından farklı değildir (bak. Problem 22 nin başlangıcı). Gerçekten, $A_k = \alpha_1, \alpha_2, \cdots, \alpha_k$ deneyinin entropisini hesaplarken her tartmanın entropisinin $\log 3$ e eşit olduğunu farz ettik, fakat şimdi gözönüne alınan halde, $N = 1/2\,(3^k - 1)$ sayısı 3 ile bölünemediğinden, daha ilk tartmanın (α_1 deneyinin) entropisi $\log 3$ den küçüktür (ilk tartmanın üç neticesi eşit olarak muhtemel değildir). $N - 1 = 3/2\,(3^{k-1} - 1)$ sayısı 3 ile bölünebildiğinden, ilk tartmada her kefeye $1/3(N - 1) = 1/2(3^{k-1} - 1)$ sayıda para koymak ve diğer $1/3(N + 2) = 1/2(3^{k-1} - 1)$ sayıda parayı bir kenara bırakmak uygundur; bu halde α_1 deneyinin üç neticesinin

$$\frac{1/3\,(N-1)}{N} = \frac{1}{3} - \frac{1}{3}N, \quad \frac{1/3\,(N-1)}{N} = \frac{1}{3} - \frac{1}{3}N,$$

$$\frac{1/3\,(N+2)}{N} = \frac{1}{3} + \frac{2}{3}N$$

ihtimalleri mümkün olduğu kadar birbirine yakın olacak, ve bu deneyin $H(\alpha_1)$ entropisi maksimum olacaktır. Fakat bu takdirde, ilk tartmadan sonra mevcut olan [ve $\log 2N - H(\alpha_1)$ e eşit bulunan] belirsizlik derecesinin ancak $k - 1$ tartmadan sonra sıfıra eşit kılınabileceği gösterilebilir. Bu, en basit olarak şu tarzda ispat edilir: ilk tartmada terazinin dengede kaldığını farz edelim; demek ki kalp para bir kenara bırakılan $1/3\,(N + 2) =$

$1/2\ (3^{k-1}+1)$ sayıda paralar içindedir; şu halde geriye β deneyi için daha 3^{k+1} sayıda mümkün netice kalır (kalp para bir kenara bırakılan paralardan biridir ve bu, diğerlerinden daha ağır veya daha hafif olabilir); hakikaten vuku bulan neticeyi belirlemek için $k-1$ tartma ile verilmeyen, $\log(3^{k+1}+1)$ e eşit bir informasyon elde etmemiz lâzımdır. Benzer tarzda, α_1 deneyinin (birinci tartmanın) seçimi ne olursa olsun, bunun neticesi, β nın neticesini belirlemesi için $k-1$ sayıda tartmanın kâfi gelemiyeceği gösterilir (bak. $N=13$ hali, sahife 121).

Şu halde

$$N=\frac{1}{2}\ (3^k-1)$$

ise k tartma işlemi kâfi gelmeyebilir. Şimdi $N\leq 1/2(3^k-3)$ ise [yani $k\geq \log_3(2N+3=\log(2N+3)/\log 3$ ise] k tartma işleminin kâfi geleceğini gösterelim *.

En basit probelm ile başlayalım: aralarından biri kalp olan N tane para gözönüne alalım; diğer taraftan ayrı olarak, kalp olmadığı bilinen bir para verilmiş olsun; kalp paranın bulunması ve bunun diğerlerinden ağır veya hafif olduğunun belirlenmesi isteniyor. Bu halde, evvelki halde olduğu gibi, $N>1/2\ (3^k-1)$ ise, deneyin neticesini belirlemek için k sayıda tartma kâfi gelmeyecektir (kalp olmadığı bilinen bir paranın ilâvesi, tabiatiyle, deneyin belirsizlik derecesini azaltmaz). Buna karşılık, artık $N=1/2\ (3^k-1)$ ise k tartmanın kâfi gelmeyeceğinden

*) Bu meselenin aşikâr olan iki istisnai hali vardır: $N=1$ ise tabiatiyle kalp paranın diğerlerinden ağır veya hafif olduğu belirlenemez; $N=2$ ise kalp parayı belirlemek imkânsızdır.

emin değiliz. Gerçekten, yardımcı parayı kullanarak, birinci deneyin üç neticesi için birbirlerine daha yakın ihtimaller elde edilebilir, ve bunun neticesi olarak evvelkine nazaran daha büyük bir informasyon bulunmuş olur; bunun için, terazinin kefelerinden her birine, $1/3(N+2)=1/2(3^{k-1}-1)$ sayıda para koymak (bu $3^{k-1}-1$ sayıda paralardan birinin kalp olmadığı bilinmektedir) ve $1/3\,(N-1)=1/2\,(3^{k-1}-1)$ sayıda şüpheli parayı tartı dışı bırakmak lâzımdır. Bu takdirde, ilk deneyin değişik neticelerinin ihtimalleri kolayca görüleceği gibi,

$$\left[\frac{N+2}{3}+\left(\frac{N+2}{3}-1\right)\right]/2N=$$

$$=\frac{1}{3}+\frac{1}{6N}\cdot\frac{1}{3}+\frac{1}{6N}$$

ve

$$\frac{N-1}{3}/N=\frac{1}{1}-\frac{1}{3N}$$

dir; gerçekten bu ihtimaller evvelkine nispeten birbirlerine daha yakın olacaktır. Şu halde α_1 deneyinin $H(\alpha_1)$ entropisi, evvelkine nazaran daha büyük olacaktır. Bu küçük fark, kat'i olarak, kalp parayı ve ağırlığını k tartma ile belirlemeyi mümkün kılmak için kâfidir.

Bunu göstermek için rekürans yolu ile ispat kullanmak elverişlidir. $k=1$ (yani $N=1$ için) iddiamız aşikârdır. İddiamızın k nın belli bir değeri için doğru olduğunu kabul edelim ve, $1/2(3^k-1)<N<1/2(3^{k+1}-1)$ ise, $k+1$ sayıda tartma işleminin kâfi olduğunu gösterelim; bu takdirde idiamız tekmil hallerde doğru olacaktır. Terazi

kefelerinden birine N sayıda olan şüpheli paralardan x sayıda koyalım, diğer kefeye, $x-1$ sayıda şüpheli paralardan ve bir tane de kalp olmadığı bilinen para koyalım. Şu halde $N_1 = N - 2x - 1$ sayıda kullanılmayan para kalacaktır. x i

$$2x - 1 \leq 3^k \quad \text{ve} \quad N - 12x - 1 \leq \frac{3^k - 1}{2},$$

yani

$$3^k \geq 2x - 1 \geq N - \frac{3^k - 1}{2}$$

olacak tarzda seçelim; bunun $N \leq 1/2(3^{k+1} - 1)$ için mümkün olduğu aşikârdır gerçekten,

$$N - \frac{1}{2}(3^k - 1) \leq \frac{1}{2}(3^{k+1} - 1) - \frac{1}{2}(3^k - 1) = 3^k$$

dir. İlk tartmada, terazi dengede kalırsa, kalp paranın kullanılmayan $N_1 \leq 1/2(3^k - 1)$ sayıda paralar arasında bulunması icap eder; elimizde kalp olmadığı bilinen bir para bulunduğundan, (yapılan hipotez mucibince) bu, k tartma ile mümkündür. Aksine olarak, kefelerden biri diğerinden daha ağır gelirse, elimizde $2x - 1 \leq 3^k$ sayıda şüpheli para kalır; üstelik, kalp para a sayıda şüpheli paralar arasında ise bunun diğerlerinden hafif olduğunu ve eğer kalp para diğer b sayıda $(a + b \leq 3^k)$ paralar arasında bulunuyorsa bunun diğerlerinden ağır olduğunu biliyoruz (eğer birinci kefe ağır geliyorsa $a = x - 1$, $b = x$ dir; eğer ikinci kefe ağır geliyorsa $a = x$, $b = x - 1$ dir; bu takdirde kalp parayı belirlemek için

evvelkine ilâveten yapılacak olan k sayıda tartma işleminin kâfi geleceğini biliyoruz (bak. yukarıda problem 21 in çözümü).

Şimdi ilk olarak gözönüne aldığımız ve içersinde bir kalp para bulunan $N \leq 1/2(3^k - 3)$ sayıda paraları yeniden düşünelim. İlk tartmada terazinin her kefesine $1/2(3^{k-1} - 1)$ sayıda para koyalım; geriye

$$N_1 = N - 2\frac{1}{2}(3^{k-1} - 1) \leq$$

$$\frac{1}{2}(3^k - 3) - (3^{k-1} - 1) = \frac{1}{2}(3^{k-1} - 1)$$

sayıda kullanılmayan para kalır *. Terazi dengede kalırsa, geri kalan $N_1 \leq 1/2(3^{k-1} - 1)$ sayıda şüpheli para ilk tartmada kullanılmayan paralardır; bu takdirde, üstelik elimizde kalp olmadığını bildiğimiz $3^{k-1} - 1$ sayıda para bulunduğundan, yukarıda ispatı yapılan iddia mucibince, kalp parayı bulmak ve ağırlığını belirlemek için $k - 1$ sayıda tartma kâfi gelir. Aksine olarak, kefelerden biri diğerinden daha ağır gelirse, $3^{k-1} - 1 < 3^{k-1}$ sayıda şüpheli para kalır ve, eğer kalp para $a = 1/2\ (3^{k-1} - 1)$ sayıda paraların arasında ise bunun diğerlerinden hafif olduğunu, ve eğer diğer $b = 1/2(3^{k-1} - 1)$ $(= a)$ paralar arasında ise diğerlerinden ağır olduğunu biliyoruz; burada da kalp parayı $k - 1$ sayıda tartma işlemi ile tayin etmek mümkündür (bak. problem 21 in çözümünün son kısmı).

---

*) $N = 1/2(3^k - 3)$ olması halinde gözönüne alınan deneyin entropisi maksimumdur ve, α_1 deneyinde (ilk tartma) β hakkında bulunan $I(\alpha_1, \beta)$ informasyonu tam $\log 3$ e eşittir.

Böylece, lüzumlu tartma sayısı hakkında yukarıda ileri sürdüğümüz iddia gösterilmiş olur.

Eğer N büyük ise,

$$k-1 < \frac{\log (2N+3)}{\log 3} \leq k$$

eşitsizlikleri ile tarif edilen k sayısı yerine, çok küçük bir hata ile, $\log 2N/\log 3$ oranı alınabilir [başka ifade ile N büyüdüğü vakit, $k/(\log 2N/\log 3)$ oranı sür'atle 1 e yaklaşır].

Problemin bu nevi şartları tabiatiyle birçok tarzda değiştirilebilir. Şimdiye kadar ancak bir kalp para olduğunu (daha doğrusu, ağırlığı diğerlerinden farklı olan yalnız bir para olduğunu) farz ettik; tabiatiyle iki (yahut daha fazla) paranın kalp olduğunu farz edebiliriz; en güç problemler, kalp para sayısının evvelden bilinmeyen haller teşkil eder *. Kalp paraların farklı ağırlıkta oldukları yahut tekmil paraların farklı ağırlıkta oldukları da düşünülebilir. (Genel olarak çok karışık olan) tekmil bu problemlerin anahtarı informasyon teorisidir.

Yukarıdaki izahatımızı aydınlatmak için iki basit misal vereceğiz.

---

*) *Bulletin of the American Mathematical* adlı Amerikan matematik dergisinin 1956 yılında yayınlanan bir sayısında, n tane ($n > 1$) yahut bilinmeyen sayıda kalp para olduğuna göre (tekmil kalp paralar aynı ağırlıkta, diğerlerinden daha hafif veya daha ağır olmak üzere), N sayıda para arasında kalp paraları kat'î olarak bulmak için tartma işlemlerinin minimum sayısını belirleme problemi incelenmiş ve bu problemin pratik tatbikatı olabileceğine işaret edilmiştir.

PROBLEM 24. — 2 si kalp olan 7 tane madenî para düşünelim; aynı ağırlıkta olan bu iki kalp para diğerlerinden ağır olsun. Kalp paraları bulmak için, bir terazi ile ağırlık kullanmadan yapılması gereken tartma işlemlerinin minimum k sayısı nedir?

Burada β sayısının, eşit olarak muhtemel $C_7^2 = 21$ sayıda neticesi vardır (başka ifade ile, 7 para arasından 2 para C_7^2 tarzda seçilebilir); şu halde $H(\beta)$ entropisi (yani elde edilmek istenilen informasyon) $\log 21$ e eşittir; $A_k = \alpha_1, \alpha_2, \cdots, \alpha_k$ karma deneyinin entropisi ise $k \log 3$ den küçüktür. Şu halde A_k nın neticesi β nın neticesini tamamen belirlerse (yani A_k kalp paraları belirlemeyi mümkün kılarsa)

$$k \log 3 \geq \log 21, \quad 3^k \geq 21$$

yahut k tam sayı olduğundan

$$k \geq 3$$

olur.

Şimdi üç tartma işlemi ile kalp paraların nasıl bulunabileceğini gösterelim. İlk tartmada, kefelerden her birine bir, iki yahut üçer tane para konabilir; bunlara tekabül eden deneyleri sıra ile $\alpha_1^{(1)}$, $\alpha_1^{(2)}$, $\alpha_1^{(3)}$ ile göstereceğiz. Eğer p_i ile $(i = 1, 2, 3)$ $\alpha_1^{(i)}$ in E neticesinin D, G nin ifade ettikleri mâna için sahife 113), D ve G neticelerinin ihtimalleri $1/2(1 - p_i)$ ye eşit olacaktır; buradan

$$H(\alpha_1^{(i)}) = -2\frac{1}{2}(1-p_i)\log\frac{1}{2}(1-p_i) - p_i \log p_i$$

neticesi çıkarılır.

Eğer terazinin kefelerine birer kalp para, yahut birer kalp olmayan para koyarsak $\alpha_1^{(1)}$ deneyi E neticesini verecektir; 7 para içinden mümkün olan tekmil tarzlarda çift para seçme sayısı $C_7^2 = 21$ olduğundan ve çift kalp para sayısı 1 e eşit bulunduğundan, kalp olmayan paralardan teşkil edilen çift para sayısı $C_5^2 = 10$ dır ve,

$$p_1 = \frac{1+10}{21} = \frac{11}{21}, \quad \frac{1-p_1}{2} = \frac{5}{21}$$

ve nihayet

$$H(\alpha_1^{(1)}) = -\frac{10}{21}\log\frac{5}{21} - \frac{11}{21}\log\frac{11}{21} \cong 0.443$$

bulunur.

Teraziye konan dört para da kalp olmadığı yahut, kefelerden herbirine konan paralardan biri kalp diğeri kalp olmadığı takdirde $\alpha_1^{(2)}$ deneyinin neticesi E olacaktır. Birinci olayın ihtimali $C_5^4/C_7^4 = 1/7$ ye eşittir (yedi para arasından dört parayı C_5^4 tarzda almak mümkündür); mümkün hale gelince, geri kalan üç paranın kalp olmamaları ihtimali $C_5^3/C_7^3 = 2/7$ dir (ikisi kalp olan) dört parayı, her çift paranın biri kalp olmak üzere, iki çifte ayırma ihtimali $2.2/C_4^2 = 2/3$ e eşittir (dört para ile C_4^2 sayıda çift teşkil edilebilir; bunlardan 2.2 tanesi bir kalp paradan ve bir de kalp olmayan paradan teşkil olunacaktır). Buradan

$$p_2 = \frac{1}{7} + \frac{2}{7}\cdot\frac{2}{3} = \frac{1}{3}, \quad \frac{1-p_2}{2} = \frac{1}{3}$$

ve

$$H(\alpha_1^{(2)}) = -3 \cdot \frac{1}{3} \log \frac{1}{3} = \log 3 \cong 0{,}477$$

bulunur.

Nihayet, tartıya katılmayan elde kalan tek para kalp değilse (bu olayın ihtimali 5/7 ye eşittir) ve eğer üstelik, her kefede bir kalp para ve iki kalp olmayan para varsa, (böyle bir ayırma ile karşılaşma ihtimali $C_4^2 \cdot C_2^1 / C_6^3 = 3/5$ dir, zira altı para ile herbiri üç paradan teşekkül etmiş C_6^3 sayıda gurup teşkil olunabilir ve bunlardan $C_4^2 \cdot C_2^1$ tanesi bir kalp para ve iki kalp olmayan paradan teşekkül etmiş olacaktır), $\alpha_1^{(3)}$ deneyinin neticesi E olacaktır. Şu halde

$$p_3 = \frac{5}{7} \cdot \frac{3}{5} = \frac{3}{7} \quad \frac{1-p_3}{2} = \frac{2}{7}$$

ve

$$H(\alpha_1^{(3)}) = -\frac{4}{7} \log \frac{2}{7} - \frac{3}{7} \log \frac{3}{7} \cong 0.470$$

elde edilir. Şu halde $\alpha_1^{(1)}$, $\alpha_1^{(2)}$, $\alpha_1^{(3)}$ deneylerinden entropisi en büyük olanı ikincisidir (bu deneyin üç neticesi eşit olarak muhtemeldir!); bunun neticesi olarak, ilk tartmada terazinin kefelerinin herbirine ikişer para koymak en elverişli olur. Muhtelif mümkün halleri inceleyeceğiz.

A — $\alpha_1^{(2)}$ deneyinin neticesi E olsun; bu takdirde iki kalp para tartıya katılmamıştır, yahut herbiri terazi

kefelerinden birindedir. Tekmil değişik hallerin incelenmesi üzerinde, informasyon teorisi bakımından durmayacağız (bak. problem 22 nin çözümü); sadece, çözümü elde etmek için kâfi derecede elverişli olan bir hali gözönüne alacağız*. Bu, terazide (meselâ sağ kefede) bulunan paralardan biri yerine, tartıya katılmayan paralardan birini koymaktır (α_2 deneyi)**. Bu şartlar altında:

A_1 — Eğer α_2 deneyinin neticesi E ise, kalp paraların ya ikisi de tartıya katılmayan paralardır, yahut biri «solun» iki parasından biri, ve diğeri, sağ kefede birinci ve ikinci tartmada bulunan paradır. Bu iki halden hangisi ile karşı karşıya bulunulduğunu belirlemek için, «solun» iki parasının ağırlıklarını karşılaştırmak kâfidir (üçüncü tartma); bunların ikisi aynı ağırlıkta ise, kalp paralar tartıya katılmayan paralardır; aksi halde, «solun» iki parasından hafif olanı ile, «sağın» iki defa tartılan parasıdır.

A_2 — Eğer α_2 deneyinin neticesi D ise, sağdaki kefeden kaldırılan para muhakkak kalptır ve, soldaki iki paradan hangisinin kalp olduğunu belirlemek için bunların ağırlıklarını karşılaştırmak kâfidir (üçüncü tartı).

A_3 — Eğer α_2 deneyinin neticesi G ise, teraziye son konan para muhakkak kalptır; henüz hiçbir tartmaya

---

*) Burada informasyon teorisinin faydasının, *ilk* tartmanın nasıl yapılacağını bize göstermesi olduğuna işaret edelim; problemin geri kalan kısmı zaten daha az karışıktır.

**) Burada tablatiyle

$$H(\alpha_2) = -\frac{2}{7}\log\frac{2}{7} - \frac{2}{7}\log\frac{2}{7} - \frac{3}{7}\log\frac{3}{7} \cong 0{,}470$$

dır.

katılmayan iki paradan hangisinin kalp olduğunu belirlemek için, bunların ağırlıklarını karşılaştırmak kâfidir (üçüncü tartma).

B — Eğer $\alpha_1^{(2)}$ deneyinin neticesi D ise, kalp paralar, ya «solun» iki parası, yahut «solun» bir parası ile tartmaya katılmayan paraların biridir. Bu takdirde, «solun» iki parası aynı ağırlıkta ise (bu ise ikinci tartma ile belirlenir) her iki para da kalptır, üçüncü tartmaya lüzum kalmaz; bu iki paranın ağırlıkları farklı ise, ikisinden hafif olanı kalptır; ikinci kalp parayı bulmak için, tartıya katılmayan üç paradan herhangi ikisinin ağırlıklarını karşılaştırmak kâfidir (üçüncü tartma); bu takdirde kalp para ikisinden ağır olanıdır; eğer ikisi aynı ağırlıkta ise kalp para üçüncüsüdür.

$\alpha_1^{(2)}$ deneyinin neticesi G olması hali aynı tarzda incelenir.

PROBLEM 25. — Aynı görünüşte 5 tane madenî para gözönüne alınıyor; bunlardan (hepsi olmamak şartiyle) birkaçı kalp olup diğerlerinden daha hafiftir (kalp paraların hepsi aynı ağırlıktadır). Bir terazi vasıtası ile, ağırlık kullanmadan, kalp paraları kat'î olarak bulmak için, tartma işlemlerinin minimum sayısı nedir?

Bu halde, mümkün neticelerin toplam sayısı

$$C_5^0 + C_5^1 + C_5^2 + C_5^3 + C_5^4 = 1+5+10+10+5 = 31$$

dir ($C_5^0 = 1$, paralardan hiçbirisinin kalp olmaması neticesine, C_5^1, bir kalp para olması neticesine, C_5^2, iki kalp para olması neticesine, v.s. tekabül eder). Tekmil bu neticeler eşit olarak muhtemel oldukları gözönüne alınırsa

$$H(\beta) = \log 31$$

elde edilir. Diğer taraftan, k sayıda müteakip tartma işleminden ibaret olan $A_k = \alpha_1, \alpha_2, \cdots, \alpha_k$ deneyinin vukuundan elde edilen informasyon $k \log 3$ den küçüktür. Eğer A_k nın neticesi β deneyinin neticesi ile belirlenmiş ise,

$$k \log 3 \geq \log 31 \ , \ 3^k > 31 \ ,$$

ve k tam sayı olduğundan

$$k \geq 4$$

bulunur.

Şu halde kalp paraları dört tartmadan daha az sayıda tartma ile bulmak her vakit mümkün değildir. Diğer taraftan, $\log 3^4 = \log 81$, $H(\beta) = \log 31$ den çok büyük olduğundan, dört tartmanın kolayca neticeyi vereceği ümit edilebilir [gerçekten, bu takdirde, α_i $(i = 1, 2, 3, 4)$ tartmalarından herbirinin entropisi log 3 ten hayli küçüktür]; hattâ, ilk tartmanın iki değişik halinin herbirine bağlı entropiyi hesap etmek lüzumsuzdur. Hakikatte, iki değişik hal, problemi çözmeyi mümkün kılar.

A. — İlk tartmada terazi kefelerinin her birine, *yalnız bir* para konduğunu farz edelim. Bu takdirde kefelerden biri diğerinden ağır gelirse, bu iki paradan hafif geleni kalp paradır. Şimdi diğer iki paranın ağırlığını ilk ikisinin ağırlığı ile karşılaştıralım; eğer daha ağır gelirlerse bunlar kalp değildir; eğer daha hafif iseler ikisi de kalptır; iki halde de son beşinci paranın kalp olup olmadığını belirlemek kalır, bu ise bu paranın ağırlığının diğer dört paradan birinin ağırlığı ile karşılaştırılarak yapılır

(üçüncü tartma); bu halde dördüncü tartmaya lüzum kalmaz. İkinci para içfti ilki ile aynı ağırlıkta ise, bunun da biri kalp diğeri kalp değildir; kalp olanını belirlemek için, bunların ağırlıklarını karşılaştırmak kâfidir (üçüncü tartma); bu halde sonuncu paranın kalp olup olmadığını belirlemek kalır, bu ise bir dördüncü tartma ile yapılır.

Şimdi ilk iki paranın aynı ağırlıkta olduğunu farz edelim; bu takdirde ikisi de kalptır, yahut ikisi de kalp değildir. Bunların ağırlıklarını ikinci para çiftinin ağırlığı ile karşılaştıralım. Bu ikinci para çifti ilki ile aynı ağırlıkta ise, sonuncu paranın ağırlığı, diğer dört paradan birinin ağırlığı ile karşılaştırılır (üçüncü tartma; bu halde dördüncü tartmaya lüzum kalmaz); eğer beşinci para diğer dört para ile aynı ağırlıkta ise, hiç kalp para yoktur; eğer daha hafif ise, kalp para yalnız budur; eğer daha ağır ise, ilk dört para kalptır. Eğer ilk para çifti ikinciden daha hafif ise, ilk iki para kalptır; ikinci çiftin paralarının ağırlıkları karşılaştırılarak (üçüncü tartma), bunların arasında kalp para olup olmadığını belirliyeceğiz; nihayet beşinci paranın kalp olup olmadığını belirlemek kalır, bu ise bir dördüncü tartma ile yapılır. İkinci para çiftinin birinciden daha hafif olması halinin incelenmesini okuyucuya bırakıyoruz.

B. — İlk tartmada terazi kefelerinden herbirine *iki* para konduğunu fark edelim. Her halde, müteakip tartmalarda, «solun» iki parası, sonra «sağın» iki parası karşılaştırılarak (ikinci ve üçüncü tartmalar; zaten birçok hallerde üçüncü tartmaya lüzum kalmaz), ilk dört para arasında hangilerinin diğerlerinden daha hafif olduğunu belirleyebiliriz. Nihayet dördüncü tartma ile beşinci paranın ağırlığı belirlenir.

İnformasyon teorisi, bilhassa, (para sayısı) N büyük bir sayı olması halinde kalp paraları bulmak için lüzumlu tartma sayısı olan k nın büyüklük mertebesini belirlemek istenildiği vakit faydalıdır Meselâ, N tane para arasında ikisinin kalp ve diğerlerinden hafif olması halinde (bak. problem 24), basit bir informasyon hesabı, N büyük ise k nın

$$\frac{\log C_N^2}{\log 3} \cong \frac{\log 1/2\, N^2}{\log 3} = \frac{2 \log N}{\log 3} - 0{,}63 \ldots$$

mertebesinde olduğunu düşünmemizi mümkün kılar (herhalde bu, $C_N^2/\log 3$ den küçük olmayacaktır); üstelik, daima ilk tartmada kefelerden herbirine iki para koymak daima elverişli olacaktır. Eğer N sayıda paranın bazıları kalp olup diğerlerinden hafif ise (bak. problem 25) ve eğer N büyük ise, k,

$$\frac{\log (C_N^0 + C_N^1 + \cdots + C_N^{N-1})}{\log 3} =$$

$$= \frac{\log (2N-1)}{\log 3} - \frac{\log 2N}{\log 3} - 0{,}63\, N$$

oranının büyüklük mertebesinde olacaktır (herhalde bundan küçük olmayacaktır), v.s.

BÖLÜM IV

HAT YOLU İLE ULAŞTIRMA PROBLEMLERİNE İNFORMASYON TEORİSİNİN TATBİKİ

§ 1. Temel Anlamlar. Bir Şifrenin Ekonomisi

Bölüm II de izah ettiğimiz entropi ve informasyon anlamlarının daha iyi anlaşılabilmesi için Bölüm III de, matematik oyunları nev'inden bazı «eğlenceli» problemleri inceledik. Bu problemlerin «oyun» olma karakteri bazı okuyucuların ilgisini azaltabilir; halbuki, informasyon teorisinin başında bu teorinin çok mühim pratik tatbikatı olduğunu söylemiştik. Bu bölümde, daha evvel izahını verdiğimiz analmları basit fakat çok ciddî problemlere, hat yolu ile haber ulaştırma problemlerine tatbik edeceğiz. Bu problemlerin, «düşünülen» bir sayının, yahut muayyen bir sayıda tartma işleminden sonra kalp paraların belirlenmesi gibi, «eğlenceli» problemlerle birçok müşterek noktaları bulunacaktır; bu sebepten, evvelki paragraflarda kullanılan birçok ispatlardan ulaştırmanın teknik problemlerinde faydalanılacaktır.

Evvelâ bir informasyonun hat yolu ile ulaştırılmasının şemasını inceleyelim; misal olarak telgrafı ele alalım. Hattın bir ucundan, «gönderen», lâtin alfabesinin 26 harfi, yahut başka bir alfabenin 32 harfi veya 10 rakkam (şifreli ulaştırma) veyahut aynı zamanda harfler ve rakkamlar yardımiyle bir *haber* gönderir. Âdi hatlı telgraf halinde bu işareti ulaştırmak için telgrafçının, karakteristiklerini

isteğine göre değiştirebileceği bir devamlı akım bulunur; böylece telgrafçı seri halinde işaretler gönderir ve bunlar, hattın diğer ucunda bulunan ikinci telgrafçı tarafından alınır. Umumiyetle pratikte kullanılan en basit işaretler *bir akımın gönderilmesi* (yani belirli bir zaman için bir devirin kapatılması) ve *poz* denilen akımın yokluğu (yani belirli zaman için devirin açılması) dır; belirli işaretler yerine belirli harfler ve rakkamlar konulması önceden tespit edilmesi şartı ile, herhangi bir haber akım ve poz göndermeleri ile ulaştırılabilir.

Ulaştırmalarda, her ulaştırılan habere tekabül eden kaide muhtelif işaretlerin bir terkibi olup buna *şifre* denir (meselâ telgraf halinde bu, telgraf şifresidir); bir haberi muhtelif işaretlere çevirme işlemine haberi *şifreleme* denir. Sadece iki basit işaretle teşkil edilen şifrelere (meselâ bir akımın ve bir pozun gönderilmesi gibi) *ikili şifreler* denir; üç işaretten faydalanan şifrelere *üçlü şifreler* denir, v.s. Telgrafta çok kullanılan Morse şifresinde («Morse Alfabesi»), her harf veya her rakkam, çok kısa akım göndermeleri («noktalar») ve üç defa daha uzun süren («çizgi») akım göndermeleri ile gösterilir, «nokta» ve «çizgiler» in araları, «nokta» süresinde, çok kısa pozlarla ayrılmıştır; harfler arasında beyaz boşluk, üç zaman birimine eşit (yani toplam süresi bir «çizgi» ninkine eşit olan müteakip üç poz süresi kadar) bir zaman kadar akım kesmesi ile, ve kelimeler arasında beyaz boşluk ise, altı zaman birimine eşit bir zaman kadar akım kesimi ile gösterilmiştir. Teleskriptör âletlerinde çok kullanılan Baudot şifresinde her harfe beş elementer işaretin muayyen bir terkibi (yani eşit sürede akım ve işaretlerin gönderilmesi) tekabül eder. Birinci işaret için iki imkânı, ikinci işaret için iki imkânla sonra, üçüncü işaret için iki imkânla ve sonra dördüncü

işaret için iki imkânla ve nihayet beşinci işaret için iki imkânla tertipleyerek $2^5 = 32$ değişik kombinezon elde edilir; şu halde Baudot şifresi, elementer şekli ile, farklı 32 harfi ulaştırmayı mümkün kılar *. Bazı telgraf âletlerinde yalnız akım kesilmeyip akımın yönü de değiştirilebilir; bu takdirde, akımı kesmek ve göndermek yerine, elementer işaretler olarak, akım bir yahut diğer yönde gönderilir; akımı bir yahut diğer yönde göndermek ve akımı kesmek suretiyle üç imkândan da faydalanılabilir. Elementer işaretlerin sayısı arttırılırsa daha kısa bir şifre yapmak (başka ifade ile, verilen bir işareti ulaştırmak için lüzumlu elementer işaretlerin sayısını azaltmak), aynı sürede elementer işaretler yardımı ile daha çok önemli sayıda «harfler» ulaştırmak mümkün olur. Daha çok karışık telgraf âletleri de vardır; bunlarda akım sadece yön bakımından değil şiddet bakımından da değişebilir; bu takdirde daha da büyük sayıda elementer işaretler elde etmek mümkündür. Telsiz telgrafta, akımın şiddetinin değişim-

---

*) Şu noktaya işaret edelim ki, Baudot şifresinde olduğu gibi, tekmil harfler, hepsi aynı sürede olan elementer işaretlerin kombinezonları ile ulaştırılmışlardır (böyle bir şifreye *düzgün* şifre denir); böyle bir şifrede iki işareti özel bir işaretle ayırmak lüzumsuzdur; zira, tekmil beş işaretli guruplar, yeni bir harfin ulaştırılması olduğu malûmdur (alıcı âlette, bu beş elementer işaretli guruplara ayırma otomatik olarak yapılabilir).

32 kombinezon, tekmil harfleri ve tekmil rakkamları ulaştırmak için kâfi gelmediğinden, Baudot şifresi üzerinde çalışan âletler, iki alıcı ile mücehhezdirler; alıcının değiştirilmesiyle aynı bir işaretler kombinezonu, diğer bir işareti ulaştırmak için kullanılabilir. Şu halde mümkün hallerin sayısı iki kat olur ve bu ise, tekmil harfleri, tekmil rakkamları ve tekmil noktalama işaretlerini ulaştırmayı mümkün kılar. Yalnız bir alıcı kullanılırsa, her harfe yahut her rakkama bir altı elementer işaret kombinezonu tekabül ettirmek mümkündür, böyle şifreler bazen telgrafta kullanılmıştır.

leri yerine, (yüksek frekansta bir elektrik salınımı olan) bir radyo dalgasının muayyen parametrelerinin değişimi alınmıştır; şu halde elementer işaretler diğer bir mânayı haizdir; bununla beraber, burada da her harfe, alıcı âlet tarafından alınan, bir takım elementer işaretler tekabül ettirilir. Hatlarla ulaştırma sistemlerinin çoğunda durum bunun benzeridir *.

Şimdi teknik mülâhazalardan sıyrılarak, ulaştırma tekniğinde rastlanan problemin genel tarzda bir matematik şeklini verebiliriz. n harfli bir alfabe (meselâ 27 harfli lâtin alfabesi, yahut 10 rakkamdan veya 37 harf ve rakkamdan veyahut harflerden, rakkamlardan ve noktalama işaretlerinden teşekkül etmiş bir alfabe) ile yazılmış bir haber olsun. Bahis konusu olan bu haberi şifrelemek, başka ifade ile, bu habere, ulaştırma «alfabesi» ni teşkil eden m sayıda elementer işaretler arasından alınmış bir takım işaretler tekabül ettirmektir; bu, en elverişli olarak nasıl yapılır?

İlk yapılacak şey, burada kullanılan «elverişli» kelimesinin mânasını tarif etmektir. *Bir haberi ne kadar az elementer işaretlerle ulaştırmak mümkün ise, bu elementer işaretlerin ait olduğu şifreyi o kadar elverişli olarak*

*) Televizyon, telefon ve radyotelefonun prensip hakkından telgraftan farklı oldukları düşünülebilir; gerçekten bunlarda, sonlu sayıda harfler yardımı ile şifrelenmiş «haberler» yerine, devamlı olarak değişen büyüklükler ulaştırılır, bunlar, bir televizyon ekranının devamlı parlaklığı, yahut bir sesin yahut bir muzik parçasının değişen yükseklik derecesi yahut genliğidir. Hakikatte, gözümüz ve kulağımız, ekrana çok yakın sahaları ve çok yakın aralıklarla birbirini takip eden sesleri ayırt edemezler; keza birbirine çok yakın iki parlaklığı ve iki genliği ayırt edemezler. Burada da, şu netice çıkarılır ki ulaştırılmış haber, muayyen sayıda aralıklı işaretlerden teşekkül etmiş olarak düşünülebilir ve bu işaretler, belirli bir parlaklık ve genlik derecesine malik olacaktır.

gözönüne alacağız. Eğer tekmil elementer işaretlerin aynı süre devam ettiklerini düşünürsek, en elverişli şifre, ulaştırmayı en çabuk mümkün kılan şifre olacaktır.

Düşüncelerimizi daha açık ifade etmek için $n = 10$ ve $m = 2$ halini inceliyelim (bu, meselâ, şifrelenmiş bir haberin akım ve poz ile gönderilmesi halidir). Burada en basit şifre, tekmil rakkamları *ikili sistem* de yazmaktan, yani 0 ı 0 ile, 1 i 1 ile, 2 yi 10 ile, 3 ü 11 ile, 4 ü 100 ile, 5 i 101 ile, 6 yı 110 ile, 7 yi 111 ile, 8 i 1000 ile, 9 u 1001 ile göstermekten ibarettir *. Sonra, dört rakkamdan daha az rakkamı olan sayılar, başına konmuş sıfırlara tekabül eden bir sayı ile tamamlanır (başka ifade ile, 0 yerine

---

*) Bütün dünyaca kullanılan ondalık sistemde her sayı 10 un kuvvetlerinin bir toplamı oolarak gösterilmiştir:

$$a = a_k \cdot 10^k + a_{k-1} \cdot 10^{k-1} + a_{k-2} \cdot 10^{k-2} + \ldots + a_1 \cdot 10 + a_0,$$

burada a_k, a_{k-1}, a_{k-2}, ..., a_1, a_0 sayının rakkamları olup 0 ile 9 arasında değerler alabilirler; a sayısı bu rakkamalrı sıralamak suretiyle yani a_k, a_{k-1}, a_{k-2}, ..., a_1, a_0 şeklinde gösterilir. Benzer tarzda, ikili sistemde, her sayı 2 nin kuvvetlerinin bir toplamı ile gösterilmiştir:

$$b = b_k \cdot 2^k + b_{k-1} \cdot 2^{k-1} + \ldots + b_1 \cdot 2 + b_0;$$

fakat burada b_k, b_{k-1}, ... b_1, b_0 rakkamlarının hepsi 2 den küçüktür, başka ifade ile, bunlar ancak 1 ve 0 değerlerini alabilirler. Meselâ 6 ve 9 sayıları

$$6 = 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 \ , \quad 9 = 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$$

yazılır ve şu halde bu sayılar ikili sistemde, sıra ile 110 ve 1001 ile gösterilir.

Keza, bir sayıyı tabanı m olan bir sistemde yazmak, bu sayıyı m nin kuvvetlerinin bir toplamı şeklinde ifade etmek demektir; bu sayıyı teşkil eden «rakkamlar» ancak 0, 1, 2, ..., $m-1$ değerlerini alabilirler.

0000, 10 yerine 0010, 101 yerine 0101 yazılır; eğer bu yapılmasa idi, Morse alfabesinde olduğu gibi, bu rakkamların herbirinden sonra bir özel işaret göndermek icap edecekti). Her rakkamın yerine 0 yahut 1 rakkamlarından dördünü koymakla bir ikili şifre teşkil edilir (üstelik bu, bir *düzgün* şifredir; bak. sahife 138 deki not). İkili şifrede yazımlış bu haberi ulaştırmak için, tabiatiyle, her 1 rakkamı için bir akım ve her 0 rakkamı için bir poz göndereceğiz.

Bir rakkamı ulaştırmak için lüzumlu ikili işaret sayısı probleminin, 0 ile 9 arasında «düşünülen» bir tam sayıyı bulmak için (bak. problem 19) sorulması icap eden (ve sadece «evet» ve «hayır» ile cevap verilen) soru sayısı problemi ile tamamen aynı olduğuna işaret edelim. Hakikatte, (0 ile 9 arasında bulunan her sayıya 0 ve 1 rakkamlarının bir kombinezonuna tekabül ettirmeyi mümkün kılar) bir ikili şifrede yazılmış herhangi bir sayı verildiğine göre, bu sayıyı bulmak istiyorsak, ilk soruda, şifrelenmiş olan sayının ilk rakkamının 1 olup olmadığını, ikinci soruda ikinci rakkamının 1 olup olmadığını v.s. sorabiliriz. Sayıyı şifrede yazmak için lüzumlu ikili işaret sayısı kadar soru sormak icap edecektir. Karşıt olarak, 0 ile 9 arasında, «düşünülen» bir tam sayıyı bulmak için lüzumlu soru dizisi bilindiğine göre, bu on sayının herbirine 0 ve 1 rakkamlarından ibaret olan bir dizi tekabül ettirebiliriz şöyle ki, ilk sorunun cevabı «evet» ise bu dizinin ilk rakkamı 1, «hayır» ise 0 dır, keza dizinin ikinci rakkamı ikinci sorunun cevabı ile belirlenir v.s.; bu tarzda bir soru dizisi bir ikili şifre tarif eder.

İkili numaralama sisteminin, bizi, problem 19 da tavsiye olunan soru serisine sevk ettiğini işaret edelim.

Bu noktayı açıklamak için, aranılan sayının 2 nin bir kuvvetine eşit miktarda değerler alabileceğini farz edelim; başka ifade ile, bu sayının 10 değer değil, fakat meselâ $2^4 = 16$ kadar (0 ilâ 15 arasında) değerler aldığını farz edelim; ikili sistemde 10 sayısı 1010, 11 sayısı 1011, 12 sayısı 1100, 13 sayısı 1101, 14 sayısı 1110 ve nihayet 15 sayısı 1111 yazılacaktır. Şu halde, ilk soru: «ikili şifrenin ilk rakkamı 1 midir?» demek, «düşünülen sayı 7 den büyük mü?» demektir, zira yalnız 8 ilâ 15 arasındaki sayıların ikili sistemdeki karşılıklarının ilk rakkamı 1 dir. İkinci soru: «ikinci rakkamı 1 e eşit mi?», ilk soruya müspet cevap halinde, «düşünülen sayı 11 den büyük mü?» demektir. Zira (ikili sistemde 7 den büyük sayılar arasında ikinci rakkamı 1 e eşit olanlar yalnız 12, 13, 14 ve 15 dir); ilk soruya bir menfi cevap halinde, bu, «düşünülen sayı 3 den büyük mü?» demektir (zira, 7 den küçük veya buna eşit sayılardan, yalnız 4, 5, 6 ve 7 nin ikili sistemde ikinci rakkamları 1 dir); v.s. Bu gösterdiğimiz husus ikili numaralama sisteminin mânasını açıklar.

Şimdi gözönüne alınan ikili şifrenin *verimliliği* meselesine geçelim. (Onlu sistemde yazılmış) şifreli haberin her rakkamı 10 mümkün değer alabilir; başka ifade ile, bunun ihtiva edebileceği informasyon $\log 10 = 1$ dir; bu informasyon miktarı ancak, tekmil rakkamlar kendi aralarında müstakil iseler, ve eğer herbiri aynı ihtimal ile 10 mümkün değeri alabilirse, elde edilecektir. Şifrelenmiş haberin her rakkamı (yani ulaştırılan her elementer işaret: akım yahut poz gönderilmesi) iki mümkün hale malik olabilir ve şu halde en çok $\log 2$ ye eşit (yani bir ikili birim) bir informasyon verir. I. Bölüm 1. paragrafının başında bir onlu birimin yaklaşık olarak 10/3 ikili birime (daha doğru olarak $1/\log 2 = 3,3219 \ldots$ ikil birime)

eşit olduğunu söylemiştik. Halbuki, gözönüne alınan şifre ile bir sayıyı ulaştırmak için dört elementer işarete, ve N sayılı bir haberi ulaştırmak için $4N$ elementer işarete lüzum vardır. $4N$ tane elementer işaretten teşekkül etmiş bir haberi ulaştırırken $4N$ tane ikili birime eşit informasyon elde edilebilecek ve bu, N onlu sayılı bir haberdeki maksimum informasyondan yaklaşık olarak $2/3N$ ikili birim kadar fazla olacaktır (bu son informasyon maksimum N tane onlu birime eşittir). Şu halde gözönüne alınan şifrenin en ekonomik olmadığı aşikârdır. Aynı haberi ulaştırmak için, ulaştırılan elementer işaretlerin sayısını azaltmak mümkündür. Gözönüne alınan şifrenin neden en ekonomik olmadığını anlamak kolaydır: gerçekten, şifrelenen haberin tekmil rakkamları müstakil değillerdir ve bunlar 0 ve 1 değerlerini aynı ihtimal ile almazlar (meselâ, ilk haberde, 0 ile 9 arasındaki sayılar eşit olarak muhtemel ise, şifrelenmiş haberde 0 rakkamını ulaştırma ihtimali, 1 rakkamını ulaştırma ihtimalinden $25/15 = 5/3$ defa daha büyük olacaktır)*. Bununla beraber ulaştırılan 0 ve 1 dizisinin maksimum informasyon ihtiva etmesi için bu dizinin tekmil rakkamlarının iki mümkün değeri aynı ihtimal ile almaları gerekir (ve üstelik, bunlar müstakil olmalıdır).

Şimdi bir ikili şifrenin nasıl inşa edileceğini gösterelim. Ulaştırılacak haberi iki rakkam gurubuna ayıralım ** ve ikili şifre vasıtasiyle, her sayıyı değil, fakat iki sayı gurubunu ulaştıralım. Böyle bir ikili sayı gurubunu (bun-

*) Gerçekten, ikili dört işaret vasıtasiyle 0 ile 9 arasındaki tekmil yazıların yazılmasında 0 rakkamına 25 defa ve 1 rakkamına sadece 15 defa rastlanacağını sağlamak kolaydır.

**) Bu iki sayı guruplarına ayırma, tabiatiyle, haberin yüz taban sisteminde şifrelenmesine denktir.

lar 00 ve 99 arasındadır), şifrelemek için lüzumlu ikili işaret sayısı 100 den küçük olan «düşünülen» bir sayıyı bulmak için sorulması gereken soru sayısına eşittir; şu halde bu 7 ye eşittir (bak. problem 19). Bu tarzda iki rakkamlı bir sayı gurubunu ulaştırmak için, (evvelce olduğu gibi $2.4 = 8$ değil) 7 elementer işaret lâzımdır; başka ifade ile N sayıyı ulaştırmak için (basitleştirmek için N yi çift farzediyoruz), $3{,}5\,N$ elementer işaret lâzımdır, bu ise ilk şifrelemeye nisbeten $N/2$ işaret noksandır. Çok sayı ulaştırılacak ise (yani N büyük ise fayda o nisbette fazla olur).

Haberi üç sayılı guruplarla kısımlara bölmek ve elde edilen her kısmı ikili sistemde şifrelemek daha da elverişli olur. Üç sayılı bir gurupu ulaştırmak için 10 elementer işaret lâzımdır. (Bak. promleb 19); şu halde bu usul ile N sayıda bir haber $10\,N/3$ elementer işaret ile ulaştırılabilecektir. Bu kısımlardan herbirini ikili sisteme çevirmeden evvel, haberi daha çok sayılı guruplarda kısımlara ayırma pratikte pek elverişli olmaz (üç sayılı kısımlardan dörtlü guruplara geçerken, şifrelemenin daha az elverişli olduğu da vâkidir; gerçekten, dört sayılı bir kısmı ulaştırmak için $14 = 3{,}5 \cdot 4$ elementer işaret lâzım olduğunu görmek kolaydır); bununla beraber, kâfi derecede büyük kısımlara ayırarak şifreyi daha da «daraltmak» ve böylece limit değerine istenildiği kadar yaklaşmak mümkün olur. [N sayıda bir haberi ulaştırmak için gerekli elementer işaret sayısı, $\log 10/\log 2)\ N = 3{,}3219 \ldots N$ dir; bak. problem 19].

Böylece, informasyona ait elementer mülâhazalarla, haberi önemli kısımlara bölerek ve bu kısımlardan her birini ikili sisteme çevirerek elde edilen şifreleme, sayılardan herbirini doğrudan doğruya ikili sisteme (yahut iki

sayılı kısımlarda ikili sisteme) çevirmekten daha ekonomiktir. Bu M sayıda kısımlara ayırarak şifreleme usulü, hakikatte, M tane işaret ile yazılan bir sayının alabileceği tekmil mümkün değerleri (burada baştarafa bir yahut daha fazla sıfır konularak yazılanlar da dahildir), sayı itibariyle eşit iki bölüme ayırmak ve, gözönüne alınan kısım birinci bölüme ait ise bir poz («0 rakkamı») ulaştırmak, ikinciye ait ise bir akım («1 rakkamı») göndermekle aynıdır; bundan sonra, gözönüne alınan kısmın ait olduğu bölüm sayı itibariyle eşit iki kısma bölünür ve yeniden, gözönüne alınan kısım bu yeni bölümlerin birine yahut diğerine ait olduğuna göre, bir poz veya bir akım ulaştırılır, v.s. Bu takdirde, M büyük ise ve ilk haberi teşkil eden sayılar müstakil ve 0 ile 9 arasında tekmil değerleri aynı ihtimal ile alabilirler ise ikili şifrede yazılmış tekmil rakkamların iyi bir yaklaşımla aralarında müstakil olacakları ve, 0 ve 1 değerlerini aynı ihtimal ile alacakları sağlanabilir. Bu takdirde, basit bir informasyon hesabı ikili şifreleme için lüzumlu rakkam sayısının

$$\frac{\log 10}{\log 2} M = \frac{M}{\log 2}$$

ye çok yakın olduğunu derhal gösterir.

Yukarıdaki mülâhazalar tabiatiyle m sayıda elementer işaret kullanılması haline teşmil olunabilir. Bu halde, M sayılı kısımları, ikili sisteme değil, fakat tabanı m olan sisteme çevirmek lâzımdır. Eğer m, 10 un bir kuvvetine (meselâ $m = 10^p$) eşit ise, M nin değeri için p ile bölünebilen herhangi bir sayı almak (özel olarak $m = 10$ ise, haberi teşkil eden sayıların herbiri ayrı ayrı ulaştırılabilir) lâzımdır, ve bu takdirde M sayıda işaretli bir kısmı

ulaştırmak için lüzumlu elementer işaret sayısı M/p ye eşit olacaktır. Eğer m sayısı 10 un tam bir kuvvetine eşit değilse, bu elementer işaretlerin sayısı, $(\log 10/\log m)M = M/\log m$ kesrini ihtiva eden en küçük tam sayıya eşit olacaktır; en ekonomik şifreleme, haberi büyük kısımlara ayıran olacaktır; bu takdirde N gibi büyük sayıda bir haberi ulaştırmak için lüzumlu elementer işaret sayısı $N/\log m$ ye çok yakın olacaktır. Eğer $m = 3$ ise, elde edilen neticeler, bir kalp parayı bulmak için lüzumlu tartma sayısını aradığımız vakit, Bölüm III ün 2. paragrafında elde edilen neticelere yakın olacaktır: her tartmanın üç neticesi olabileceğinden, müteakip tartmaların neticeleri üç değer alabilen * bir rakkam serisi ile, yani üçlü taban sisteminde yazılmış bir sayı şeklinde gösterilebilir.

Neticelerimizi, ulaştırılacak haberin, «rakkamlı» olarak değil de, n «harfli» bir «alfabe» vasıtasiyle (meselâ, lâtin alfabesi, rus alfabesi, yahut bir alfabenin harfleri ve rakkamlar vs. vasıtasiyle) şifrelenmiş olması halinde de kolayca alabiliriz. Bu halde, alfabenin tekmil «harfler» ini 0 ile $n-1$ arasındaki rakkamlarla numaralamak kâfidir; bu takdirde haber n tabanlı bir sistemde yazılmış bir «sayı» şeklini alır. Bu haberi, herbiri m değer alabilen bir elementer işaretler serisi vasıtası ile ulaştırmak için, bunu (ya tamamen yahut kısımlara ayırarak) bir m tabanlı sisteme çevirmek ve elde edilen rakkamların herbirini bir işaret ile ulaştırmak lâzımdır. Üstelik, kâfi derecede büyük kısımlara ayırarak, N harfli bir haberi ulaştırmak için lüzumlu işaret sayısının $(\log n/\log m)$ N ye istenildiği kadar yakın olması daima temin edilebilir.

*) Üçlü taban sisteminde âdet olduğu vechile bu değerler, 0, 1, 2 rakkamları ile gösterilebilir, fakat; meselâ E, D ve G ile de göstermek mümkündür.

§ 2. Shannon - Fano şifresi. Şifrelemenin temel teoremi

«Alfabe» harflerinin sayısı n ye eşit ise ve eğer elementer işaretlerin sayısı m ise, ilk bakışta, $\log n/\log m$ oranının, mümkün olan en elverişli şifreyi karakterize ettiği düşünülebilir. Hakikatte, alfabemizin N tane harfi $N \cdot \log n$ ye eşit bir informasyon ihtiva edebilir; bu informasyonu, m değer alabilen bir seri işaretle ulaştırmak için, en az $N \cdot (\log n/\log m)$ sayıda işaret ulaştırmak lâzımdır.

Bu neticenin «informasyon» tâbiri kullanılmadan elde edilebileceğine işaret edelim; bunun için sadece muhtelif «N harfli haberler» ve «N, sayıda elementer işaret serileri» hesap edilebilir. Başka ifade ile, burada elde edilen netice, Bölüm II nin 1 - 4 paragraflarında yapılan tekmil ispatlara lüzum göstermez ve şu halde, haberlerin ulaştırılmasının teknik problemi için informasyon teorisinin büyük ehemmiyeti hakkında ileri sürülen iddiaları katiyen doğrulamaz.

Hakikatte, bir harfin ulaştırılmasının ortalama olarak $\log n/\log m$ elementer işarete lüzum gösterdiğine dair netice yanlıştır, ve Bölüm II deki muhakemeler hatanın nerede bulunduğunu göstermeyi mümkün kılar. n harfi bir alfabede N harfli bir haberin en fazla $N \cdot \log n$ ye eşit bir informasyon ihtiva ettiği tabiatiyle doğrudur; fakat, hakikatte, böyle bir metin, eğer bunun bir mânası var ise (başka ifade ile mevcut bir dilde yazılmış ise), *asla böyle bir informasyon ihtiva etmez.* Bu, her harf ancak eşit olarak muhtemel olduğu, başka ifade ile alfabenin her harfine aynı sıklıkla rastlanabildiği takdirde maksimum

log n informasyonunu ihtiva edebileceğinden, aşikârdır (bak. Bölüm II), fakat herhangi bir türkçe metinde a ve e harflerine c ve f harflerine nazaran daha pek çok rastlanır; böyle intizamsızlıklar tekmil dillerde bulunur. Şu halde bir harfte bulunan informasyon daima log n den hayli küçüktür, bu ise, bize, böyle bir metini ulaştırmak için daha çok elverişli bir şifre kullanma imkânını verir.

Bölüm II de tesis edilmiş formüller, makul bir tarzda yazılmış bir metinin bir harfinde bulunan ortalama informasyonun log n den (n, alfabenin harflerinin sayısıdır) ne dereceye kadar küçük olduğunu değerlendirmeyi mümkün kılar. Bir türkçe metin gözönüne alalım. Türkçe alfabede 29 harf vardır; şu halde bir harfte bulunan informasyon (yani ulaştırılan bir harfin alınmasından ibaret olan deneyin vuku bulmasında elde edilen informasyon) en fazla

$$H_0 = \log 29 \cong 1{,}462 \quad \text{(onlu birimlerle)}$$

e eşit olacaktır.

Fakat bu informasyonun elde edilebilmesi için alfabenin tekmil harflerine aynı sıklıkla rastlanmalıdır; halbuki kâfi derecede uzun bir metinin incelenmesi farklı harflere ait sıklıkların birbirlerinden çok farklı olduğunu gösterir. Türkçe alfabede harflere ait sıklıklar aşağıdaki cetvelde gösterilmiştir:

	Harf	1000 harfteki sıklık
1	*a*	99,92
2	*i*	93,81
3	*e*	88,13
4	*n*	80,27
5	*l*	79,09
6	*r*	70,01
7	*ı*	50,65
8	*d*	48,67
9	*k*	44,36
10	*u*	37,49
11	*y*	36,70
12	*m*	35,72
13	*t*	31,59
14	*s*	26,69
15	*b*	24,53
16	*ü*	21,59
17	*z*	20,61
18	*ş*	19,62
19	*o*	17,86
20	*ğ*	13,57
21	*v*	12,36
22	*g*	10,99
23	*h*	10,79
24	*ç*	10,00
25	*c*	8,63
26	*p*	7,26
27	*ö*	6,05
28	*f*	4,51
29	*j*	0,40

Şu halde türkçe metinin bir harfinde bulunan ortalama informasyon olarak

$$H_1 = -0{,}099 \cdot \log 0{,}099 - 0{,}093 \cdot \log 0{,}093 - \ldots$$
$$\ldots - 0{,}0004 \cdot \log 0{,}0004 \cong 1{,}328 \text{ (onlu birimlerle)}$$

elde ederiz, 0,099 ve 0,093 yukardaki cetvelin ilk iki harfine ait sıklık, 0,0004 ise cetvelin son harfine ait sıklıktır.

Yukarıdakine benzer bir hesap tekmil lisanlar için yapılabilir. Harflere ait sıklıklar gözönüne alınırsa, ortalama informasyon olarak, bir fransızca metinde onlu birimlerle $H_1 \cong 1{,}199$, bir ingilizce metinde $H_1 \cong 1{,}242$, bir almanca metinde $H_1 \cong 1{,}233$ ve bir ispanyolca metinde $H_1 \cong 1{,}209$ elde ederiz. Tekmil hallerde bu değerlerin 1,462 den hayli küçük olduklarını görüyoruz; esasen bu değerler dilden dile değişir, zira harflere ait sıklıklar tekmil dillerde aynı değidir *.

Fakat, alfabenin farklı harflerine ait farklı sıklıklar hesaba katılarak da, bir harfte bulunan informasyon için, hakikaten ulaştırılan değerden daha büyük bir değer elde edilir. Gerçekten, meselâ bir türkçe metin için onlu birimlerle 1,328 e eşit bir informasyon ancak, içinde 10 000 kâğıt bulunan ve bunların 999 unda *A* harfi, 938 inde *İ*

---

*) Rusçada 31 harf vardır ve bir harfte bulunan informasyon

$$H_0 = \log 31 = 1{,}491 \quad \text{(onlu birim)}$$

dir. Harflere ait sıklıklar gözönüne alınarak elde edilen H_1 informasyonu ise

$$H_1 = 1{,}342 \quad \text{(onlu birim)}$$

dir.

harfi, ... ve 4 ünde J harfi yazılı olan bir sandıktan bir kâğıt çekilerek metinin her harfi belirlenmiş olsa idi, elde edilecekti. Böyle bir deneyin entropisi onlu birimlerle 1,328 e eşittir; eğer N sayıda aynı tip deneyler $\alpha_1, \alpha_2, \cdots, \alpha_N$ ise (her çekilişte kâğıdın sandığa tekrar konulduğu farz edilerek)

$$H(\alpha_1 \alpha_2 \cdots \alpha_N) =$$
$$H(\alpha_1) + H(\alpha_2) + \cdots + H(\alpha_N) \cong N \cdot 1{,}328$$

yazılır, başka ifade ile, ortalama informasyon, yine onlu birimlerle 1,33 değerine yakındır. Fakat bir türkçe metinin telgrafla ulaştırılmasında, verilen bir harfin ulaştırılması ihtimali kendisinden evvelki harflere tâbi olduğu görülür; meselâ son ulaştırılan bir sesli harf ise bundan sonra ulaştırılacak harfin sessiz olması ihtimail çok artar. Bu hususlara tekmîl dillerde rastlanır; meseâ ingilizcede t harfinden sonra (*«th»* bileşimini teşkil etmek için) h harfinin gelmesi çok muhtemeldir, fransızcada q harfinden sonra büyük bir ihtimal ile u harfi gelir. Bir mânası olan bir metinde, iki müteakip harfi belirlemekten ibaret olan iki deney α_1 ve α_2 ile gösterilirse, bu iki deneyin vuku bulması ile elde edilen informasyon

$$H(\alpha_1 \alpha_2) = H(\alpha_1) + H_{\alpha_2}(\alpha_2) < H(\alpha_1) + H(\alpha_2)$$

dir. Bir önceki harfin, haberin alınması ile daima bilindiği hesaba katılırsa, ortalama informasyon şu formül ile hesaplanmalıdır:

$$H_2 = H_{a_1}(a_2) =$$

$$p(a_1a_1)\log p(a_1) - p_{a_1}(a_2a_1)\log p_{a_2}(a_1) - \cdots$$
$$\cdots - p(a_na_1)\log p_{a_n}(a_1) -$$
$$- p(a_1a_2)\log p_{a_1}(a_2) - p(a_2a_2)\log p_{a_2}(a_2) - \cdots$$
$$\cdots - p(a_na_2)\log p_{a_n}(a_2) -$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$- p(a_1a_n)\log p_{a_1}(a_n) - p(a_2a_n)\log p_{a_2}(a_n) - \cdots,$$
$$\cdots - p(a_na_n)\log p_{a_n}(a_n)$$

burada n, alfabe harflerinin sayısı, $p(a_i, a_j)$, a_j harfinin a_i harfinden $(i, j = 1, 2 \cdots, n)$ sonra gelmesi ihtimali ve $p_{a_i}(a_j) = p(a_i a_j) / p(a_i)$ ise, bir evvelki a_i harfi olduğu bilindiğine göre, a_j harfinin bulunması ihtimali (bağlı ihtimal)dir. Keza, evvelki iki harfin bilindiği gözönünde tutularak, bir harfin ulaştırılmasında bulunan ortalama informasyon şu formül ile verilmiştir:

$$H_3 = H_{a_1a_2}(a_3)$$
$$= - p(a_1a_1a_1)\log p_{a_1a_1}(a_1) - p(a_1a_2a_1)\log p_{a_1a_2}(a_1) - \cdots$$
$$\cdots - p(a_na_na_1)\log p_{a_1a_2}(a_1) -$$
$$- p(a_1a_1a_2)\log p_{a_1a_1}(a_2) - p_{a_1a_2}(a_1a_2a_2)\log p(a_2) - \cdots$$
$$\cdots - p(a_na_na_2)\log p_{a_na_n}(a_2) -$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$- p(a_1a_1a_n)\log p_{a_1a_1}(a_n) - p(a_1a_2a_n)\log p_{a_1a_2}(a_n) - \cdots$$
$$\cdots - p(a_na_na_n)\log p_{a_na_n}(a_n),$$

burada $p(a_i\, a_j\, a_k)$ ile «$a_i\, a_j\, a_k$» harf dizisinin bulunması ihtimali, $p_{a_i a_j}(a_k)$ ile de, $a_i\, a_j$ harflerinin kombinezonundan sonra a_k harfinin bulunmasının bağlı ihtimali gösterilmiştir. Çok uzun bir metinin ulaştırılması isteniyorsa, benzer bir formül yardımı ile bir harf tarafından verilen informasyonu hesap etmek lâzımdır; fakat burada büyük sayıda önceki harfleri hesaba katmak icap eder.

$$H_N = H_{\alpha_1 \alpha_2 \cdots \alpha_{N-1}}(\alpha_N) =$$

$$= -\,p(a_1 \cdots a_1\, a_1) \log p_{a_1 a_1 \cdots a_1}(a_1) -$$

$$-\,p(a_1 \cdots a_1\, a_2) \log p_{a_1 \cdots a_1}(a_2) - \cdots$$

$$\cdots -\,p(a_n \cdots a_n\, a_n) \log p_{a_n \cdots a_n}(a_n)$$

iformasyonu ancak N arttığı zaman azalır ve belirli bir H_∞ limitine yaklaşır ve bu limit, bir uzun metin ulaştırılması halinde, bir harfte bulunan informasyonun teorik ortalama değeri olarak gözönüne alınabilir.

SHANNON, ikili ve üçlü harf guruplarından hareket ederek ingilizce dili için H_2 ve H_3 informasyonlarının hesabını yapmıştır. H_2=1,075 (onlu birimler) ve H_3=0,993 (onlu birimler) bulmuştur (ingilizcede, tekmil harflerin müstakil olduklarını, fakat normal izafi sıklıklara malik bulundukları farz edilerek, bir harfte bulunan ortalama informasyon $H_1 = 1.242$ onlu birimidir). İngilizce dilinin muayyen kelimelerinin izafi sıklıklarına ait bazı istatistik verilenleri kullanarak, SHANNON bu dilde, H_8 için (yani ilk evvelki 7 harfi malûm farz ederek, bir harfte bulunan informasyonu) yaklaşık olarak 0,6 (onlu birim) bulmuştur.

Daha uzun bir harf gurupu gözönüne alınsaydı (bir metinde kelimelerin bileşimi hakkında statistik verilenler vasıtasiyle), bir harfte ortalama olarak bulunan informasyonda yeni bir azalma olacağı haklı olarak düşünülebilir.

Böylece, mânayı haiz bir metinin bir harfinde bulunan ortalama informasyonun, alfabenin bir harfinde bulunabilecek olan maksimum informasyondan iki defa daha küçük olduğunu görüyoruz (ingilizce dili için $H_0 = \log 26 \cong 1{,}415$ ve $H_8 \cong 0{,}6$ değerleri karşılaştırılabilir). SHANNON'un yaptığı gibi, bir dilde yazılmış bir metinin bir harfinde bulunan informasyonun alfabenin bir harfinde bulunan ortalama informasyona oranının birimden farkına (yani $1 - H_\infty / H_0 = 1 - H_\infty / \log n$ ifadesine) bu dilin *kalanı* der isek, Avrupa dillerinin çoğunda bu kalan % 50 mertebesinde (hattâ daha fazla) dır; başka ifade ile, herhangi bir metinde, harflerin en az % 50 si dilin kendi yapısı ile belirlenmişlerdir. Şimdi (meselâ bir ikili şifre kullanılan) bir şifreleme problemine dönersek, informasyon kemmiyetinin hesabı, (N kâfi derecede büyük olmak üzere) N harften teşekkül etmiş mânalı bir metinin bir ikili şifre vasıtasiyle, yalnız ikili şifrenin (0 ve 1 rakkamları) $H_\infty / \log 2$ sayıda elementer işaretini kullanarak, ulaştırmak imkânını ortadan kaldırmadığı neticesine varırız (burada H_∞, bir harfte bulunan ortalama informasyonun limiti olup meselâ H_{100} e yakındır) H_∞ değeri $H_0 = \log n$ den hayli küçük olduğundan elementer işaretlerin, haberimizi ulaştırmak için lüzumlu olan $N \cdot H_\infty / \log 2$ toplam sayısı, evvelce elde edilen $N \cdot \log n / \log 2$ sayısından hayli küçük olabilir. Aynı zamanda, N harften teşekkül etmiş bir metini ulaştırmak için $N \cdot H_\infty / \log 2$ den daha az ikili işaret kullanan bir şifrelemenin mümkün olmadığını görmek kolaydır: bu mümkün olsa idi; mutlaka bir

informasyon kaybı olacak, ve son haber, ilk haberin yeniden tesisini mümkün kılmıyacaktı.

Bölüm III ün 1. paragrafında problem 20 nin incelemesinde söylenenler mucibince en elverişli şifrenin elde ediliş tarzını anlamak kolaydır. Evvelâ, metini, N harfli kâfi derecede uzun kısımlara ayırmak uygun olur (bak. sahife 144-145) sonra N harfle teşkili mümkün olan tekmil kısımların tam bir listesini yapmak (eğer alfabede n harf var ise, n^N sayıda kısım olacaktır) ve bunları eksilen ihtimal sırasiyle (en fazla muhtemel olan başta, en az muhtemel olan sonda olmak üzere) tasnif etmek lâzımdır; bu ihtimallerin hesabı, harflerin herbirinin izafi sıklıkları ve her dilin hususiyetlerinden olan, harflerin guruplaşmalarının sıklıkları hesaba katılmalıdır. Sonra, bu kısımlar, birbirlerine en yakın ihtimaller gözönüne alınmak suretiyle iki gurupa ayrılır; birinci gurupa ait olan kısımlara şifrelemenin ilk rakkamı olarak 1 rakkamı tekabül ettirilir; ikinci gurupun kısımlarına ise 0 rakkamı tekabül ettirilir. Sonra iki guruptan herbiri, ihtimalleri mümkün olduğu kadar yakın iki alt-grupa ayrılır ve gözönüne alınan kısım birinci veya ikinci alt-grupa ait olduğuna göre, şifrelemenin ikinci rakkamı olarak (ve şifrelemenin ilk rakkamından müstakil olmak üzere) 1 veya 0 tekabül ettirilir. Daha sonra dört alt-grupun herbiri yine ihtimalleri mümkün olduğu kadar yakın iki alt-grupa ayrılır, vs., işleme, her alt-grupta yalnız bir kısım kalıncaya kadar devam edilir. Böylece bir şifreleme usulü ilk defa 1948-1949 da C. E. SHANNON ve R. M. FANO tarafından ortaya konmuştur; bu sebepten çoğu zaman buna SHANNON-FANO şifresi denir. Misal olarak, aşağıdaki cetvel, ihtimalleri (eksilme sırasına göre) 0,3; 0,2; 0,1 (iki kısım); 0,05; 0,03 (5 kısım); 0,02 (2 kısım); 0,01 (6 kısım) olan 18

kısmı SHANNON-FANO metoduna göre şifreleme tarzını gösterir *.

Kısım numarası	İhtimal	Şifre numarası
1	0,3	11
2	0,2	10
3	0,1	011
4	0,1	0101
5	0,05	0100
6	0,03	00111
7	0,03	00110
8	0,03	00101
9	0,03	00100
10	0,03	00011
11	0,02	000101
12	0,02	000100
13	0,01	000011
14	0,01	0000101
15	0,01	0000100
16	0,01	000001
17	0,01	0000001
18	0,01	0000000

Romen rakkamları grup ve alt-grup numaralarıdır.

*) Çok kısımlar olması halinde, ihtimalleri yakın gruplara ve alt gruplara ayırma işi ziyadesiyle sıkıntılı olabilir; bu işi basitleştirmek için mihaniki usuller mevcuttur.

Temeli SHANNON-FANO şifresi olan esas prensip, şifre numarasının her rakkamını, informasyonu maksimum olacak tarzda, yani kendisinden evvelki rakkamlardan müstakil olmak üzere 0 ve 1 değerlerini mümkün olduğu kadar birbirlerine yakın ihtimallerle alacak tarzda seçmekten ibarettir. Bu takdirde şifrenin muhtelif numaralarında rakkam sayıları tabiatiyle farklı olacaktır (yukarıda verilen misâlde bu sayılar iki ile yedi arasında değişir): SHANNON-FANO şifresi düzgün değildir. Bununla beraber, şifrelemesi daha uzun bir şifre numarasının şifrelenmesinin başlangıcı ile üstüste gelen şifre numarası olamıyacağı aşikârdır (hususi olarak bu, bu şifrenin «düşünülen» bir sayının aranması problemi ile benzerliği ile de görülür, bak. sahife 101); şu halde, bu tarzda şifrelenmiş her haber daima bir tek tarzda çözülür. Şu noktaya ehemmiyetle işaret etmek lâzımdır ki, SHANNON-FANO şifresinde, şifrenin en kısa numaraları, görülmesi ihtimali en büyük olan kısımlara tekabül eder (gerçekten, guruplara ve alt-guruplara ayırmada, en muhtemel kısımlar, yalnız bir kısım ihtiva eden alt-gruplarda daha çabuk olarak bulunur: bak. yukardaki misâl). Şifrenin bazı numaraları oldukça uzun olmakla beraber, bu uzunluğun ortalama değeri diğer her şifredekinden daha küçük olacaktır (yahut, herhalde daha büyük olmayacaktır); bu ortalama değer, şifrede informasyonun muhafazası için minimum olarak kabul edilen $H/\log 2$ değerinden pek az büyüktür. Şu halde, meselâ, yukarıda incelenen ve 18 harflik bir alfabeye tekabül eden misâlde, düzgün ikili şifrenin en iyisi, beş rakkamlı bir şifredir (gerçekten $2^4 < 18 < 2^5$ dir); şu halde bir harfi ulaştırmak için, beş elementer işaret lâzımdır; buna tekabül eden SHANNON-FANO şifresinde, ulaştırılması için yedi elementer işarete lüzum gösteren harfler vardır; bununla beraber, bir harfi

ulaştırmak için lüzumlu işaretlerin sayısının otalama değeri

$$2.0{,}5 + 3.0{,}1 + 4.0{,}15 + 5.0{,}15 + 6.0{,}06 + 7.0{,}04 = 3{,}29$$

dır. Bu değer 5 den hayli küçüktür ve

$$\frac{H}{\log 2} = \frac{-0{,}3 \log 0{,}3 - 0{,}2 \log 0{,}2 - \ldots - 6 \,.\, 0{,}01 \,.\, \log 0{,}01}{\log 2} \cong 3{,}25$$

den az farklıdır.

Gitgide daha uzun kısımları şifrelemek şartiyle, $H/\log 2$ teorik limit değerine daha da iyi yaklaşılabilir. Bu, *şifrelemenin temel teoremi denilen şu genel meselenin neticesidir: N harfli kısımlara ayrılmış bir haber şifrelendiği vakit, haberin bir harfinin ulaştırılması için lüzumlu işaretlerin ortalama sayısı* $H/\log 2$ *ye* (yani, haberin bir harfinde bulunan informasyonun bir elementer işarette bulunan maksimum informasyona oranına) *istenildiği kadar yakın olacak tarzda, N yi kâfi derecede büyük seçmek daima mümkündür.* Bu, başka tarzda da ifade olunabilir: *M harfli çok uzun bir haber, bu haberi N harfli kâfi dereecde uzun kısımlara ayırmak ve bu kısımları doğrudan doğruya şifrelemek şartiyle, sayısı* $MH/\log 2$ *ye istenildiği kadar yakın* (fakat hiçbir halde bundan küçük olmayan) *elementer işaretler vasıtasiyle şifrelenebilir.* Bu kısımların şifreleme usulünden bahsetmememiz'in bir tesadüf eseri olmadığına işaret edelim: gerçekten, ileride, mümkün olan şifreleme usullerinin çok değişik olduğunu göreceğiz (şu halde, muhakkak SHANNON-FANO usulünü aynen takip etmek lüzumlu değildir). Şu halde, en ekonomik şifrenin aranmasında temel rolü

oynayan, çok uzun kısımlara ayırmadır. Gelecek paragrafta, bu çok uzun kısımlara ayırmanın, bir bulanıklık mevcut olması halinde (bu, ulaştırma hattının normal çalışma şartlarına tekabül eder), şifreleme usulleri farklı olmasına rağmen, büyük faydaları olduğunu göreceğiz.

Şimdi temel teoremimizin ispatına geçelim; daima, haberimizin müteakip harflerinin birbirlerinden müstakil olduklarını farz edeceğiz. Aşağıda verilen ispat uzunca olmasına rağmen çok faydalı ve entropi anlamının mânasını da iyice açıklar. Bu ispat, haberimizin harflerinin izafî sıklıklarının çok farklı olması halinde bile, ihtimallerinin toplamı ihmal edilebilen bazı kısımlar hariç, tekmil kısımlara aynı uzunlukta numaralar tekabül ettiren «hemen hemen düzgün» bir şifrenin kullanılabileceğini gösterir. «Az muhtemel» kısımlara gelince bunların herhangi tarzda şifrelenebileceği kolayca anlaşılır: pratikte bunun hiçbir ehemmiyeti yoktur, zira bunların görülmesi ihtimali çok azdır.

Evvelâ, «alfabe» nin, ihtimalleri $p_1=p$ ve $p_2=1-p=q$ olan, sadece a ve b gibi iki harften teşekkül ettiği çok basit halden başlayalım. Bu takdirde, bu alfabenin N harfinden teşekkül etmiş kısımların toplam sayısı 2^N olacaktır (bak. sahife 138-139). Fakat bu kısımların çoğunun ihitmali pek küçük olacaktır: «Alfabe» nin iki harfinin izafî sıklıkları p ve q olduğundan, mühimce bir ihtimale malik olan kısımlar, N büyük ise, N_p defa a harfini ve N_q defa b harfini ihtiva eden kısımlar olacaktır. Daha kesin bir tarzda, N çok büyük ise, ε çok küçük olmak üzere (meselâ ε, 0,001, yahut 0,0001 veyahut 0,0000001 olabilir, ε için bu sayılardan herhangi birisi alınabilir, hattâ daha da küçük bir sayı alınabilir, yeter ki N çok büyük olsun) a harfinin izafî sıklığı $p-\varepsilon$ ve $p+\varepsilon$

arasında bulunmayan tekmil kısımların ihtimallerinin toplamı çok küçüktür ve bunlar ihmal edilebilir. a harfine $N(p-\varepsilon)$ ve $N(p+\varepsilon)$ arasında bulunan bir sayı defa rastlanan kısımlara gelince, bunların her birinin ihtimalleri daha da küçük olacaktır (eğer N çok büyük ise, kısımların toplam sayısı çok büyük ve her kısmın kendi ihtimali tabiatiyle çok küçük olacaktır), fakat bu ihtimallerin toplamı 1 e çok yakın bulunacaktır:

a harfine tam N_p defa * rastlanıldığı N harfli kısımların sayısı, N elemanın, N_p, N_p alınmış kombinezonuna eşittir:

$$C_N^{Np} = \frac{N!}{(Np)!(N-Np)!} = \frac{N!}{(Np)!(Nq)!}.$$

Bizi ilgilendiren hal, N, Np ve Nq sayılarının çok büyük olması halidir (bu sayılar, kısımlarımızdaki a ve b harflerinin sayısını gösterirler). O halde

$$\left(\frac{n}{e}\right)^n < n! < n\left(\frac{n}{e}\right)^n,$$

eşitsizliklerini kullanalım, burada e, 2 ile 3 arasında bir sayıdır (bak. kitabın sonunda ilâve II). Buradan

$$N! < N\left(\frac{N}{e}\right)^N, \quad (Np)! > \left(\frac{Np}{e}\right)^{Np}, \quad (Nq)! > \left(\frac{Nq}{e}\right)^{Nq}$$

---

*) Eğer Np tam sayı değilse bunun yerine buna en yakın olan K tam sayısını alacağız, N nin çok büyük olması halinde Np ile K arasındaki fark çok küçük olacaktır (bak. sahife 103). Aynı husus Nq sayısı için de vakidir.

buradan

$$C_N^{Np} < \frac{N\left(\frac{N}{e}\right)^N}{\left(\frac{Np}{e}\right)^{Np}\left(\frac{Nq}{e}\right)^{Nq}} = \frac{N\frac{N^N}{e^N}}{\frac{N^N p^{Np} q^{Nq}}{e^N}} = \frac{N}{p^{Np} q^{Nq}}$$

bulunur. Tahminen, a harfine $Np+1$, $Np+2$, $\cdots$, $Np+N\varepsilon$ defa, yahut $Np-1$, $Np-2$, $\cdots$, $Np-N\varepsilon$ defa rastlandığı kadar kısım vardır (gerçekten, tekmil hallerde, a harfinin sıklık farklarının, hesabını yaptığımız sıklık farkına nispeten çok küçüktür). Şu halde, «muhtemel» kısımların yani ihtimallerinin toplamı ihmal edilebilenlerden gayrı tekmil kısımların toplam sayısının

$$N_1 = 2N\varepsilon \cdot \frac{N}{p^{Np} \cdot q^{Nq}} = \frac{2N^2\varepsilon}{p^{Np} \cdot q^{Nq}}$$

sayısından küçük olduğunu farz ederek çok küçük bir hata yaptığımıza hükmedebiliriz. Şimdi «muhtemel» kısımlar için en ekonomik şifreyi kullanarak bir harfi ulaştırmak için lüzumlu ikili işaretlerin ortalama sayısını hesap etmek kolaydır. Meselâ N_1 tane (yahut daha az) kısmı şifrelemek için, bunları sayı bakımından eşit iki gurupa ayırırız ve kısım birinci gurupa ait ise 1 rakkamını, ikinci gurupa ait ise 0 rakkamını ulaştırırız; sonra bu guruplardan her birini sayı bakımından eşit ikişer alt-grupa ayırırız ve gözönüne aldığımız kısım, bu alt-gruplardan birine yahut diğerine ait bulunduğuna göre 1 yahut 0 rakkamını ulaştırırız vs. Bu takdirde, her kısmın ulaştırılmasında lüzumlu ikili işaret sayısı, kısımların toplam

sayısının logaritmasının $\log 2$ ye oranına eşit olacaktır (bak. problem 19 un çözümünün sonu); şu halde bu sayı

$$\frac{\log N_1}{\log 2} = \frac{\log 2\varepsilon}{\log 2} + \frac{2 \log N}{\log 2} - \frac{N(p \log p + q \log q)}{\log 2}$$

den küçük olacaktır. Eğer SHANNON-FANO metodu ile elde edilen daha elverişli bir şifre kullanılırsa, farklı kısımları ulaştırmak için ikili işaret sayıları farklı olacaktır. Bununla beraber böyle bir şifrede, bu sayılar ancak azalabilirler (bak. problem 20) *. Şu halde «muhtemel» kısımlardan birini bir SHANNON-FANO şifresi ile ulaştırmak için lüzumlu ikili işaret sayısının ortalama değerinin

$$N \cdot \frac{-p \log p - q \log q}{\log 2} + \log N \cdot \frac{2}{\log 2} + \frac{\log 2\varepsilon}{\log 2}$$

den küçük olduğunu iddia edebiliriz. Haberin bir harfini ulaştırmak için lüzumlu ikili işaretler sayısının ortalama değerine gelince, bu,

$$\frac{H}{\log 2} + \frac{\log N}{N} \cdot \frac{2}{\log 2} + \frac{1}{N} \cdot \frac{\log 2\varepsilon}{\log 2}$$

den küçüktür, burada

$$H = -p \log p - q \log q$$

---

*) Hakikatte, «muhtemel» kısımlara tatbik edilen SHANNON-FANO metodu büyük bir fayda temin etmez, zira bu kısımların ihtimalleri kâfi derecede birbirlerine yakındır (gerçekten, harflerin her birinin izafi sıklığı tekmil bu kısımlarda aşağı yukarı aynıdır).

dür. Yukarıdaki formülde, $H \to \infty$ olduğu vakit, $H/\log 2$ den sonra gelen iki terim sıfıra yaklaşır, buradan, bir SHANNON-FANO şifresi ile ulaştırılan yalnız «muhtemel» kısımlar gözönüne alınırsa, kâfi derecede uzun bir haberin bir harfini ulaştırmak için lüzumlu ikili işaretler sayısının ortalama değerinin $H/\log 2$ ye istenildiği kadar yakın olduğu neticesi çıkarılır *.

«Az muhtemel» olan kısımlara gelince, bunları şifrelemek için $H/\log 2$ den birçok defa büyük ikili işaretler sayısı kullansak bile, bir harfi ulaştırmak için lüzumlu işaretler sayısının ortalama değerine ancak pek az tesir eder, zira tekmil bu kısımların ihtimallerinin toplamı çok küçüktür.

$p_1 + p_2 + \cdots + p_n = 1$ olmak üzere ihtimalleri p_1, p_2, $\cdots$, p_n olan n harfli bir alfabe halinin incelenmesi çok uzundur. N harfli tekmil kısımlar arasında, en muhtemel olanlar tahminen ilk harfe Np_1 defa, ikinci harfe Np_2 defa $\cdots$ n yinci harfe Np_n defa rastlananlar olacaktır. İlk harfe tam Np_1 defa, ikinci harfe tam Np_2 defa, $\cdots$, n yinci harfe tam Np_n defa rastlanan kısımların sayısı, N elemanın, herbiri sıra ile Np_1, Np_2, $\cdots$, Np_n eleman ihtiva eden, n gurupa ayırma sayısına eşit olacaktır; bu sayının

$$\frac{N!}{(Np_1)!\,(Np_2)!\cdots(Np_n)!}$$

ye eşit olduğunu göstermek kolaydır (bu formül, kombinezon sayısı formülünü genelleştirir).

*) Burada, bu ortalama değerin, $H/\log 2$ den küçük olamıyacağına işaret edelim.

$$N! < N\left(\frac{N}{e}\right)^N$$

ve

$$(Np_1)! > \left(\frac{Np_1}{e}\right)^{Np_1}, \quad (Np_2)! > \left(\frac{Np_2}{e}\right)^{Np_2}, \quad \cdots,$$

$$\cdots, \quad (Np_n)! > \left(\frac{Np_n}{e}\right)$$

eitsizliklerini kullanarak, N nin kâfi derecede büyük değeri için, kısım sayısının

$$\frac{N}{p_1^{Np_1} \cdot p_2^{Np_2} \cdots p_n^{Np_n}}$$

den küçük olduğu gösterilebilir.

Ancak, ihtimali, birinci harf için $p_1 - \varepsilon$ ile $p_1 + \varepsilon$ arasında, ikinci harf için $p_2 - \varepsilon$ ile $p_2 + \varepsilon$ arasında, $\cdots$, n yinci harf için $p_n - \varepsilon$ ile $p_n + \varepsilon$ arasında olan kısımlar gözönüne alınırsa, bu kısımların sayısı için

$$(2N\varepsilon) \cdot \frac{N}{p_1^{Np_1} \cdot p_2^{Np_2} \cdots p_n^{Np_n}} = \frac{2^n \varepsilon^n N^{n+1}}{p_1^{Np_1} \cdot p_2^{Np_2} \cdots p_n^{Np_n}}$$

elde ederiz. Hiç olmazsa içindeki harfelrden biri için ihtimal, yukarıda gösterilen olmayan diğer kısımlara gelince, bunların ihtimallerinin toplamı sonsuz küçük olacaktır ve bunu hesaba katmayabiliriz.

O halde, «muhtemel» kısımları, problem 19 (sahife

101 - 104) da gösterildiği gibi şifrelesek bile, her kısmı ulaştırmak için,

$$H = -p_1 \log p_1 - p_2 \log p_2 - \cdots - p_n \log p_n$$

olmak üzere,

$$N \cdot \frac{H}{\log 2} + \log N \cdot \frac{n+1}{\log 2} + \frac{n \log 2\varepsilon}{\log 2}$$

den fazla ikili işaret kullanmayacağız; «muhtemel» kısımları şifrelemek için, (III. bölümün 1. paragrafındaki problem 20 nin çözümüne tekabül eden) SHANNON-FANO metodunu kullanırsak bu sayı gösterilen değerden *haydi haydi* küçük olacaktır. Buradan, bir harfi ulaştırmak için lüzumlu ikili işaretler sayısının ortalama değerinin

$$\frac{H}{\log 2} + \frac{\log N}{N} \cdot \frac{n+1}{\log 2} + \frac{1}{N} \cdot \frac{n \log 2\varepsilon}{\log 2}$$

den küçük olacağı neticesi çıkarılır. $N \to \infty$ için bu değer

$$\frac{H}{\log 2}$$

ye yaklaşır, bu ise, böyle bir şifreyi kullanarak bir harfi ulaştırmak için lüzumlu ikili işaretler sayısının ortalama limit değerini verir, ve göstermek istediğimiz neticedir.

Nihayet, biraz evvel verdiğimiz ispatın temelini teşkil eden esas prensip üzerinde bir kerre daha duralım. n harfli bir alfabede alınmış tekmil mümkün kısımlar (başka ifade

ile, n sayıda neticesi olan ve N defa tekrar edilen bir deneyin N sayıda tekmil neticeler serisi) gözönüne alınırsa, böyle kısımların (yahut bu serilerin) toplam sayısı

$$n^N = 10^{N \log n}$$

dir. Bu kısımların her birinin ihtimali, ve hattâ bu kısımlardan muayyen sayıdakinin ihtimallerinin toplamı, N büyük olması şartiyle, çok küçük olacaktır. En az muhtemel olan kısımlardan, ihtimallerinin toplamı kâfi derecede küçük (önceden tespit edilen bir δ sayısından daha küçük) olacak tarzda, muayyen sayıdakilerini hesaba katmazsak, δ *ne olursa olsun* (istenildiği kadar küçük), N yi kâfi derecede büyük seçmek şartiyle ve H entropiyi göstermek üzere geriye, sayısı,

$$\left(\frac{1}{p_1}\right)^{Np_1} \left(\frac{1}{p_2}\right)^{Np_2} \cdots \left(\frac{1}{p_n}\right)^{Np_n} = 10^{N.H}$$

ile aynı mertebeden olan kısımların kaldığını gösterdik. Burada hemen şunu söyleyelim ki, H, $\log n$ den daha küçük olduğundan tekmil harflerin yahut tekmil neticelerin eşit olarak muhtemel olması halini hesaba katmayarak N sonsuz olarak büyür ise, «muhtemel» kısımların sayısı, kısımların toplam sayısından daha çok küçüktür («muhtemel kısımların sayısının tekmil kısımların sayısna oranı olan $10^{HN}/10^{N \log n} = 10^{-N(\log n - H)}$, $N \to \infty$ için sür'atle sıfıra yaklaşır). Bundan başka, N kâfi derecede büyük ise, «muhtemel» kısımlarda muhtelif harfler için, $p_1, p_2, \cdots, p_n$ ye istenildiği kadar yakın izafi sıklıklar elde edilebileceğini gösterdik. Verilen bir kısmın ihtimali ancak, bu kısmın harflerinin izafi sıklığına tâbi olduğundan (birinci harfe

N_1 defa, ikinci harfe N_2 defa, $\cdots$, n yinci harfe N_n defa rastlanan bir kısmın ihtimali $p_1^{N_1} p_2^{N_2} \cdots p_n^{N_n}$ dir), N kâfi derecede büyük olması şartiyle, tekmil «muhtemel» kısımların hemen hemen aynı ihtimale malik olduklarını kolayca göstermek mümkündür. Başka ifade ile, biz burada, problem 12 nin sonunda da italik harflerle yazılı özelliği gösterdik, işte, şifreleme teorisinde entropinin oynadığı temel rolü izah eden bu özeliktir.

§ 3. Ulaştırmada bir bulanıklığın mümkün olması hali

Şimdiye kadar yaptığımız tekmil incelemelerde, ulaştırmanın hatasız, yani bulanıklık bulunmadan yapıldığı farz edildi. Pratikte asla bu böyle olmaz: ulaştırma sırasında, başka bir alanın tesiri ile elektrik alanında ve bunun neticesi olarak gönderilen işarette bir değişiklik daima mümkündür. Hususi olarak, telgrafta, iki elementer işaret, bir akım ve bir poz ulaştırıldığı vakit, alıcı istasyonun bulanıklık dolayısiyle, bir akım yerine bir poz ve bir poz yerine bir akım alması için daima muayyen bir ihtimal mevcuttur. Bulanıklığın az olması ve işaretlerin alınmasında yapılacak hatanın ihmal edilebilmesi haline evvelki mülâhaza ve ispatlarımız tatbik edilebilir. Fakat genel halde de informasyon teorisi çok faydalı bir rol oynar.

Basitleştirmek maksadiyle, hepsi birbirinden müstakil olan n harfli bir alfabe gözönüne alalım ve bu harflerin ulaştırılacak metinde herhangi bir anda ulaştırılması ihtimalleri $p_1, p_2, \cdots, p_n$ olsun. $A_1, A_2, \cdots, A_m$ gibi m sayıda elementer işaret kullanan bir ulaştırma hattı gözönüne alalım; bu hat vasıtasiyle bir zaman biriminde böyle

elementer işaretlerden L sayıda ulaştırılabildiğini farz edelim. Şu halde, yukarıda elde edilen temel neticeyi şöyle ifade edebiliriz: *bulanıklık olmaması halinde gözönüne alınan hat bir haberi*

$$v = \frac{L \log m}{H} \quad \text{(zaman biriminde harf sayısı)}$$

ye istenildiği kadar yakın bir hızla ulaştırır (burada

$$H = -p_1 \log p_1 - p_2 \log p_2 - \cdots - p_n \log p_n$$

olup bir harfin ulaştırılmasında bulunan entropiyi gösterir), fakat hiç bir halde bundan büyük olamaz. Böyle bir ulaştırma hızını elde etmek için, haberi kâfi derecede uzun parçalara ayırmak ve, bu kısımları ulaştırmak için SHANNON-FANO şifresini, başka ifade ile, şifrelenmiş haberin «kalan» ı mümkün olduğu kadar çok az olan bir şifre kullanmak kâfidir.

Bulanıklıkların bulunması muhtemel ise mesele biraz değişir. Yalnız bu halde, ulaştırılmış işaretlerden sonra bir «kalan», ulaştırılmış haberi alınmış işaretlerden hareket ederek doğru olarak yeniden tesis etmeyi mümkün kılar; mühim bir bulanıklık ihtimali varsa, meselâ her ulaştırılmış kelimeyi birçok defa tekrarlayarak yahut her harf için bir kelime (bu harfle başlayan kelimeyi) ulaştırarak, ilk haberin «kalan» ını dahi arttırmayı tecrübe ederiz. Bizi, minimum bir «kalan» a malik olan bir şifrelemeye sevk eden SHANNON-FANO şifresinin burada kullanılamıyacağı ve mutlaka ulaştırma hızının çok küçük olacağı aşikârdır. Şu halde hız ne kadar azaltılmalıdır?

Bu soruyu cevaplandırmak için evvelâ, bulanıklığın

mevcut olduğu bir hattın matematik olarak nasıl düşünülebileceğini incelemek lâzımdır. $A_1, A_2, \cdots, A_m$ gibi m tane elementer işaretle çalışan bir hatta, bulanıklık dolayısiyle, A_1 işaretini ulaştırarak A_1 işaretinin alınması ihtimali $p_{A_1}(A_1)$ (Şu halde $p_{A_1}(A_1)$, A_1 işaretini hatasız olarak alma ihtimalidir), A_1 işaretini ulaştırarak A_2 işaretini alma ihtimali $p_{A_1}(A_2)$, $\cdots$, A_1 işaretini ulaştırarak A_m işaretini alma ihtimali $p_{A_1}(A_m)$ olduğunu farz edelim. Keza A_2 işaretini ulaştırarak $A_1, A_2, \cdots, A_m$ işaretlerini alma ihtimalleri $p_{A_2}(A_1)$, $p_{A_2}(A_2)$, $\cdots$, $p_{A_2}(A_m)$; $\cdots$; A_m işaretini ulaştırarak $A_1, A_2 \cdots, A_m$ işaretlerini alma ihtimalleri $p_{A_m}(A_1)$, $p_{A_m}(A_2)$, $\cdots$, $p_{A_m}(A_m)$ olsun

$$p_{A_1}(A_1), p_{A_1}(A_2), \cdots, p_{A_1}(A_m);$$

$$p_{A_2}(A_1), p_{A_2}(A_2), \cdots, p_{A_2}(A_m);$$

$$\cdots\cdots\cdots\cdots$$

$$p_{A_m}(A_1), p_{A_m}(A_2), \cdots, p_{A_m}(A_m)$$

ihtimalleri hattımızda hasıl olabilecek bulanıklıkları istatistik tarzda gösterirler; başka ifade ile bunlar, (yapısına bağlı mümkün olan bazı bulanıklıklara malik olan) gözönüne alınan hattın bir nevi matematik mahiyetini teşkil ederler. Şimdi ulaştırılacak işaretin $p(A_1)$ ihtimali ile A_1 işareti, $p(A_2)$ ihtimali ile A_2 işareti, $\cdots$, $p(A_m)$ ihtimali ile A_m işareti olduğunu farz edelim $[p(A_1 + p(A_2) + \cdots + p(A_m) = 1$ dir$]$. Bu halde, ulaştırılmış işareti belirlemekten ibaret olan β deneyinin $H(\beta)$ entropisi

$$H(\beta) = -p(A_1) \log p(A_1) - p(A_2) \log p(A_2) - \cdots$$
$$\cdots - p(A_m) \log p(A_m)$$

ye eşittir. Hattın diğer ucundan alınmış işareti belirlemekten ibaret olan α deneyi tabiatiyle β deneyine tâbidir; α nın, β deneyinin A_j neticesinin vukuuna bağlı, A_i neticesinin $(i, j = 1, 2, \cdots, m)$ vukuu ihtimali, doğrudan doğruya $p_{ij}(A_i)$ ye eşittir. α da β hakkında bulunan ortalama informasyon

$$I(\alpha, \beta) = H(\beta) - H_\alpha(\beta)$$

dır, burada $H_\alpha(\beta)$, bir bağlı entropi olup bölüm II, paragraf 2 de verilen formüllerle hesaplanabilir (bu formüllerde k ve j yerine m, B_j yerine A_i konur: gerçekten β ve α deneyleri aynı mümkün neticelere maliktir). Tabiatiyle, $I(\alpha, \beta)$ informasyonu daima β deneyinin $H(\beta)$ entropisinden küçüktür, zira β deneyi hakkında maksimum informasyon bu deneyin vuku bulmasındaki informasyondur. $I(\alpha, \beta)$ informasyonu $H(\beta)$ entropisine, ancak, α deneyinin neticesi β deneyinin neticesini tamamen belirlemesi halinde, yani alınan işaretin ulaştırılmış olan işareti daima bir tarzda belirlemesi halinde (pratikte bu, bulanıklığın işaretlerin normal olarak alınmasına mâni olmaması demektir) eşittir; $I(\alpha, \beta)$ informasyonu ancak, α deneyi β deneyine tâbi olmaması halinde sıfıra eşittir (başka ifade ile bu, alınan işaretlerin ulaştırılan işaretlere asla tâbi olmaması halidir; bu halde artık, çok mühim bulanıklık sebebi ile, hiçbir ulaştırma mümkün değildir).

Şimdi, bir ulaştırma hattının *kapasite* anlamını izah edelim. Bulanıklık yoksa, m sayıda elementer işaretle çalışan bir ulaştırma hattının C kapasitesi

$$C = L \log m$$

formülü ile tarif edilmiştir, burada L, bu hat vasıtasiyle birim zamanda ulaştırılması mümkün olan elementer işaretlerin sayısıdır. Şu halde C kemmiyetinin, bu hat vasıtasiyle zaman biriminde ulaştırılabilen informasyon birimlerinin sayısına eşit olduğu görülür; gerçekten, her elementer işarette $\log m$ sayıda informasyon birimi bulunduğunu biliyoruz. Bulanıklık bulunabilmesi halinde, alıcı hattın alıcı ucunda bir elementer işaretten elde edilen informasyon kemmiyeti,

$$I(\alpha, \beta) = H(\beta) - H_\alpha(\beta)$$

ya eşittir; bu kemmiyet tabiatiyle, $A_1, A_2, \cdots, A_m$ işaretlerinin ulaştırılması ihtimalleri olan $p(A_1), p(A_2), \cdots, p(A_m)$ ye tâbidir. Şu halde hattın kapasitesi

$$C = L \cdot \text{maks } [H(\beta) - H_\alpha(\beta)]$$

formülü ile tarif edilmiştir; burada maks $[H(\beta) - H_\alpha(\beta)]$ ifadesi, $I(\alpha, \beta) = H(\beta) - H_\alpha(\beta)$ informasyonunun, $p(A_1)$, $p(A_2)$, $\cdots$, $p(A_m)$ değiştirildiği vakit, maksimum değeridir. C kemmiyeti tabiatiyle yine, hat vasıtasiyle birim zamanda ulaştırılabilen maksimum informasyon birim sayısına eşittir. Bu tarif mucibince, hattın kapasitesinin ancak $p_{A_1}(A_1), \cdots, p_{A_1}(A_m)$; $\cdots$; $p_{A_m}(A_1), \cdots, p_{A_m}(A_m)$ bağlı ihtimallerine tâbi olduğu aşikârdır. $k = i$ ise $p_{A_i}(A_k) = 1$ ve $k \neq i$ ise $p_{A_i}(A_k) = 0$ olması halinde, başka ifade ile ulaştırılmış işaretin hakikaten alınması (bulanıklığın olmaması) halinde

$$H_\alpha(\beta) = 0 \text{ ve maks } [H(\beta) - H_\alpha(\beta)] = \\ = \text{maks } H(\beta) = \log m$$

elde edilir (bu maksimum değer tekmil ulaştırılmış değerlerin eşit olarak muhtemel olması, yani $p(A_1) = p(A_2) = \cdots = p(A_m) = 1/m$ olması halinde elde edilmiştir; böylece, bulanıklığın mevcut olmaması halinde bir hattın kapasitesi için verdiğimiz tarif, bulanıklığın mevcut olmasının mümkün bulunduğu haldeki daha genel tarifin bir özel halidir. Genel halde, $p(A_1)$, $p(A_2)$, $\cdots$, $p(A_m)$ ihtimalleri farklı olduğu vakit, $H(\beta) - H_\alpha(\beta)$ maksimum değerini alabilir; bununla beraber, bu maksimum değerin hiçbir halde bu ihtimallere tekabül eden $H(\beta)$ entropisinden daha büyük olamıyacağı ve, $H(\beta)$ entropisinin maksimum değeri olan $\log m$ den *haydi haydi* daha büyük olamıyacağı aşikârdır; buradan anlaşılır ki, bulanıklığı mevcut olabilen bir hattın kapasitesi, aynı sayıda elementer işaret kullanan ve birim zamanda aynı sayıda elementer işaret ulaştıran bulanıklığı bulunmayan bir hattın kapasitesinden daha küçüktür.

Şu halde, bulanıklığın mevcut olmaması halinde, haberlerin ulaştırma hızına ait temel netice, hattın kapasitesi ithal olunarak şöyle ifade edilebilir: haberi kâfi derecede büyük kısımlara ayırarak ve uygun bir şifre (SHANNON-FANO şifresi) kullanarak, haberlerin ulaştırma hızı,

$$v = \frac{C}{H} \quad \text{(harf birim zamanda)}$$

ye istenildiği kadar yakın kılınabilir, fakat hiçbir vakit daha büyük olamaz (burada H, ulaştırılmış haberin bir harbinin entropisidir). SHANNON, bir ulaştırmanın, bulanıklık bulunması halinde benzer durum arz ettiğini ispat etmiştir. Daha kesin bir ifade ile, *bir hatta bulanıklık*

mevcut olabilmesi halinde uygun bir şifre yardımı ile haberleri

$$v = \frac{C}{H} \quad \text{(harf birim zamanda)}$$

ye istenildiği kadar yakın (fakat mutlaka bundan küçük) *bir hızla ulaştırmak daima mümkündür, şöyle ki, ulaştırılan her harfin* (ve bunun neticesi olarak muayyen sayıda harften teşekkül etmiş her kısmın) *alınmasında bir hata bulunması ihtimali,* önceden verilmiş ve *istenildiği kadar küçük bir* ε *sayısından* (meselâ 0.001, yahut 0.0001, veyahut 0.0000001 den) *daha küçük olsun; aksine olarak ulaştırma, birim zamanda* C/H *harften büyük bir hızla yapılmış ise, ulaştırılmış bir harfin alınmasında bir hata bulunması ihtimali* (ulaştırma hızına tâbi olan ve bununla artan) *sonlu ve belirli bir sayıdan daima daha büyük olacaktır.* Çok önemli olan bu özelliğin ispatını, ziyadesiyle karışık olduğu için, verememek zorundayız. Burada sadece bu çok mühim teoremin bir özel haline temas eden bazı hususları göstermekle iktifa edeceğiz. İki elementer işaret (meselâ bir A_1 akımı ve bir A_2 pozu) kullanan bir ulaştırma hattı gözönüne alalım ve, bu işaretlerin herbirinin hatasız ulaştırma ihtimalinin $1-\Delta$, hatalı bir ulaştırma ihtimalinin de Δ olduğunu farz edelim. Başka ifade ile,

$$p_{A_1}(A_1) = p_{A_2}(A_2) = 1-\Delta \ , \quad p_{A_1}(A_2) = p_{A_2}(A_1) = \Delta$$

farz edelim, böylece sahife 169 da verilen ihtimal cetveli burada şu şekli alır:

$1-\Delta$	Δ
Δ	$1-\Delta$

Şu halde

$$H_{A_1}(\alpha) = H_{A_2}(\alpha) = -\Delta \log \Delta - (1-\Delta) \log (1-\Delta)$$

ve

$$H_\beta(\alpha) = p(A_1) H_{A_1}(\alpha) + p(A_2) H_{A_2}(\alpha) = \\ = -\Delta \log \Delta - (1-\Delta) \log (1-\Delta)$$

dır [gerçekten, daima $p(A_1) + p(A_2) = 1$ dir]; böylece incelenmiş olan misâlde $H_\beta(\alpha)$ hiç bir şekilde $p(A_1)$ ve $p(A_2)$ ihtimallerine tâbi değildir. Fakat,

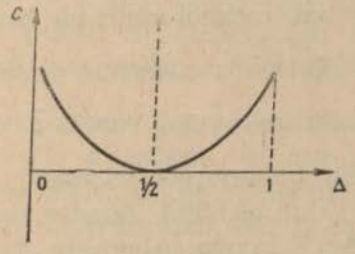
$$\text{maks } H(\alpha) = \text{maks } [-p(A_1) \log p(A_1) - \\ - p(A_2) \log p(A_2)] = \log 2$$

olduğunu biliyoruz [bu maksimum, $p(A_1) = p(A_2) = 1/2$ için elde edilmitir]; şu halde hattın kapasitesi

$$C = L \text{ maks } [H(\alpha) - H_\beta(\alpha)] = \\ = L [\log 2 + \Delta \log \Delta + (1-\Delta) \log (1-\Delta)]$$

ile verilmiştir. Bu kapasite, ulaştırmada bir hatanın Δ ihtimaline tâbidir. $C(\Delta)$ fonksiyonunun gösterdiği eğri, şekil 8 de verilmiştir. Kapasite, $\Delta = 0$ ve $\Delta = 1$ için $L \log 2$ ye eşit bir maksimum değer alır ($\Delta = 0$ bulanıklığın bulunmaması hali, $\Delta = 1$ ise her A_1 işaretine A_2 işareti ve her A_2 işaretini A_1 işaretine çeviren bulanıklık halidir, bu ikinci haldeki bulanıklığın haberin anlaşılmasına asla mani olmayacağı aşikârdır). Bundan başka, $\Delta > 1/2$ ise, alınan haberde daima, her A_1 işareti yerine A_2, ve her

A_2 işareti yerine A_1 konulabilir; bu takdirde hata ihtimali $1-\Delta < 1/2$ ye eşit olan bir ulaştırma hattı elde edilecektir. Demek oluyor ki, Δ yerine $1-\Delta$ alınırsa hattın kapasite değeri değişmez; şu halde $C(\Delta)$ fonksiyonunun gösterdiği eğri, $\Delta = 1/2$ doğrusuna nazaran simetrik olacaktır. $\Delta = 1/2$ için hattın kapasitesi sıfıra eşittir; bu halde, *hakikaten ulaştırılmış işaretten müstakil olarak*, hattın alıcı ucundan, 1/2 ye eşit bir ihtimal ile bir A_1



Şekil: 8

işareti yerine bir A_2 işareti ve, yine 1/2 ye eşit bir ihtimal ile bir A_2 işareti yerine bir A_1 işareti alınıyor demektir; şu halde, böyle bir hat üzerinde hiçbir ulaştırma mümkün değildir. İleride $\Delta < 1/2$ farz edeceğiz, zira yalnız bu hal pratik bir fayda arz eder.

Şimdi SHANNON teoremine geçelim. Bu teoreme göre, ulaştırmayı $v = \frac{C}{H} = L\frac{c}{H}$ (zaman biriminde harf) ya istenildiği kadar yakın (fakat bundan daima küçük) bir hızla yapılmayı mümkün kılacak bir şifrenin bulunması mümkündür, şöyle ki, alınan haberin çözülmesinde yapılan hata ihtimali çok küçük (daha kesin bir ifade ile, önceden verilen her küçük sayıdan daha küçük) olsun; burada

$$c = \log 2 + \Delta \log \Delta + (1 - \Delta) \log (1 - \Delta)$$

dır. Zaman biriminde L harf ulaştırabildiğimizden, bu ulaştırma hızını elde etmek için N harften teşekkül etmiş uzun bir haberin şifrelenmiş şeklinin «ortalama olarak» tahminen $(H/c)N$ (yahut çok az daha fazla) sayıda elementer işaret ihtiva etmesi lâzımdır; o halde, uzun bir T zaman aralığı zarfında LT sayıda elementer işaret, yani vT harfli bir haber ulaştıracağız.

Şimdi, N harften teşekkül etmiş bir haberin şifrelenmiş şekli $(H/\bar{c}) \cdot N$ sayıda elementer işaret ihtiva eden bir şifreleme metodu arayacağız; burada $\bar{c}$ ile, *yalnız* $\bar{c} < c$ *şartını gerçekleyen* bir sayı gösterilmiştir ($\bar{c}$ sayısı c ye istenildiği kadar yakın olabilir). Bundan başka, *tekmil* N harfli $n^N = 10^{\log n \cdot N}$ sayıda haberlerin (burada n, alfabenin harf sayısıdır) şifrelenmiş şeklinin bu unzunlukta olmasının faydasız olduğunu biliyoruz: bunun 10^{HN} sayıda «muhtemel» haberler için doğru olması kâfidir; $10^{\log n \cdot N} - 10^{N \cdot H}$ sayıda diğerlerine gelince, N çok büyük olduğu takdirde, bunların görülme ihtimallerinin toplamı çok küçüktür, ve hattâ bunların şifrelenmiş şekilleri daha çok uzun olsa bile, bu keyfiyet ulaştırma hızını azaltmayacaktır [ve bu hız, zaman biriminde $L(\bar{c}/H)$ harfe yakın kalacaktır]. Şu nokatya da işaret edelim ki, ulaştırmada çok büyük bir doğruluk elde etmek için, 10^{NH} sayıda «muhtemel» haberlerin alınmasında yapılan çözüm hatasının çok küçük olduğundan emin bulunmak kâfidir, zira tekmil diğer haberlere çok seyrek rastlanır.

Elementer işaret sayısı $(H/\bar{c}) \cdot N$ olan muhtelif kısım-

ların sayısı $2^{HN}/\bar{c} = 10^{HN \log 2}/\bar{c}$ ye eşittir; $\bar{c} < c \leq \log 2$ olduğundan bu sayı 10^{NH} dan büyüktür; şu halde 10^{NH} sayıda «muhtemel» kısımların herbirine, şifrelenmiş yazı olarak bu kısımlardan birini tekabül ettirebiliriz. Fakat, bu kısımların ulaştırılmasına ait hatanın da küçük olması lâzımdır. Mümkün olan hatalar hususunda şüpheye düşmeyeceğimiz yegâne hal, $\Delta = 0$ halidir; bu takdirde $c = \log 2$ olup $(H/c) \cdot N$ sayıda elementer işaretli tekmil kısımlar sayısı N harfli «muhtemel» haberlerin sayısına eşittir; demek oluyor ki, bulanıklığın olmaması halinde yukarıda verdiğimiz neticeyi bulacağız. Buna karşılık, eğer $\Delta > 0$ ve $c < \log 2$ (ve $\bar{c} < c$) ise, $(H/c) \cdot N$ sayıda elementer işaretli kısımların toplam sayısı 10^{HN} den büyüktür, bu ise beklenilen bir neticedir, zira 10^{HN} sayıda şifrelenmiş yazı, ulaştırmadan ileri gelen değişiklik ihtimaline rağmen tefrik edilebilmeleri için, kâfi derecede farklı olmalıdırlar.

10^{HN} sayıda şifrelenmiş yazıların çözülmesinde yapılan hata ihtimalinin küçük olması için bunların hangi şartları sağlaması gerektiği kolayca görülür. Bir elementer işaretin ulaştırılmasındaki hata ihtimali Δ olduğundan, $k = (H/c) \cdot N$ sayıda elementer işaretli bir kısmın ulaştırılmasında, bunlardan tahminen $k \cdot \Delta$ kadarı hatalı olacaktır; hayli çok yahut hayli az hata ile yapılan ulaştırma halleri, k büyük ise (yani N büyük ise), çok az muhtemeldir. Şu halde, k elementer işaretli her kısma, bu kısımdan ancak tahminen $k\Delta$ işaret kadar farklı olan, bir takım kısımlar tekabül ettirebiliriz (bu fark için daha kesin olarak, δ küçük bir sayı olmak üzere $k(\Delta + \delta)$ ile $k(\Delta - \delta)$ arasında bulunur diyebiliriz); eğer alınan haber bu kısımlara ait bulunuyorsa, ulaştırılan kısmın gözönüne alınan kısım olduğunu söyleyebiliriz. Şu halde şifre şöyle

olmalıdır: *N sayıda muhtemel harflerin 10^{HN} sayıda şifrelenmiş şekilleri böyle bir kısım takımına tekabül etmeli ve bu takımlar farklı olmalıdır;* eğer bir kısım aynı zamanda bu takımlardan birçoğuna ait olsaydı, ulaştırmanın alınmasında hangi kısmın ulaştırılmasının istenildiği belirlenemiyecek ve şu halde ulaştırılan haber çözülemiyecekti.

Şimdi her takımın sayısının $k(\Delta + \delta)$ ile $k(\Delta - \delta)$ arasında b harfini [bunun neticesi olarak $k(1 - \Delta - \varepsilon)$ ile $k(1 - \Delta + \varepsilon)$ arasına a harfini] ihtiva eden k sayıda a ve b harfli kısımların sayısına eşit olduğuna işaret edelim; gerçekten, ilk kısmı teşkil eden elementer işaretler ile üstüste gelen elementer işaretleri a ile ve, üstüste gelmeyenleri b ile gösterebiliriz. Daha önce k sayıda a ve b harfli kısımların sayısını hesaplamıştık (bak. sahife 160-161); bu sayı

$$k_1 = \frac{2k^2\delta}{\Delta^{k\Delta} \cdot (1-\Delta)^{k(1-\Delta)}} = \frac{\left(2\frac{H}{\bar{c}}N\right)^2 \delta}{\Delta^{\left(\frac{H}{\bar{c}}N\right)\Delta} \cdot (1-\Delta)^{\left(\frac{H}{\bar{c}}N\right)(1-\Delta)}}$$

sayısından küçüktür. Böylece, $(H/\bar{c}) \cdot N$ elementer işaretten teşekkül etmiş 10^{HN} sayıda şifreli yazıya, her biri en fazla k_1 kısım ihtiva eden 10^{HN} sayıda kısım takımı tekabül eder. Şu halde tekmil takımlardaki kısımların toplam sayısı

$$N_1 = 10^{HN} \cdot k_1 = 10^{HN} \frac{2\left(\frac{H}{\bar{c}}N\right)^2 \delta}{\Delta^{\left(\frac{H}{\bar{c}}N\right)\Delta} \cdot (1-\Delta)^{\left(\frac{H}{\bar{c}}N\right)(1-\Delta)}}.$$

dan küçüktür. Şimdi

$$\bar{c} < c = \log 2 + \Delta \log \Delta + (1-\Delta) \log (1-\Delta)$$

ise N_1 sayısının, elementer işaret sayısı $k = (H/\bar{c})N$ olan tekmil kısımların toplam sayısı olan

$$M = 2^k = 2^{\frac{H}{\bar{c}}N} = 10^{\log 2 \frac{H}{\bar{c}}N}$$

den küçük olduğunu gösterelim. Bunun için

$$\frac{\log N_1}{\log M} = \frac{HN + \log 2\,\delta + 2\log\left(\frac{H}{\bar{c}}N\right) - \frac{H}{\bar{c}}N[\Delta \log \Delta + (1-\Delta)\log(1-\Delta)]}{\log 2 \cdot \frac{H}{\bar{c}}N}$$

$$= \frac{\bar{c} - \Delta \log \Delta - (1-\Delta)\log(1-\Delta)}{\log 2} + \frac{\log 2\,\delta}{\log 2 \cdot \left(\frac{H}{\bar{c}}N\right)} + \frac{2\log\left(\frac{H}{\bar{c}} \cdot N\right)}{\log 2 \cdot \left(\frac{H}{\bar{c}}N\right)}.$$

oranını teşkil edelim. Yukardaki bağıntının sağ tarafında bulunan toplamın ilk terimi, hipotezlerimiz mucibince, 1 den küçük, ve diğer iki terim N büyük olduğu vakit çok küçük olduğundan (bak. sahife 162), N kâfi derecede büyük alındığı vakit yukarıdaki oranın 1 den küçük olduğundan emin olabiliriz, bu ise istenilen ispattır. Hatta, N nin kâfi derecede büyük değerleri için N_1/M oranının *çok küçük* olduğunu da iddia edebiliriz. Gerçekten, c_1 ile, c

ve $\bar{c}$ arasında bulunan herhangi bir sayı gösterilirse $(\bar{c} < c_1 < c)$, $\log N_1/\log M$ oranı

$$\lambda = \frac{c_1 - \Delta\log\Delta - (1-\Delta)\log(1-\Delta)}{\log 2} < \frac{c - \Delta\log\Delta - (1-\Delta)\log(1-\Delta)}{\log 2} = 1$$

den daha küçüktür (gerçekten daha yukarıda gözönüne alınan toplamın birinci terimi λ dan küçük ve diğer iki terimi çok küçüktür); buradan anlaşılır ki,

$$\log N_1 < \lambda \log M, \quad N_1 < M$$

ve bunun neticesi olarak,

$$\frac{N_1}{M} < \frac{1}{M^{1-\lambda}}$$

olduğu anlaşılır; $\lambda < 1$ olmak üzere

$$\frac{1}{M^{1-\lambda}} = \frac{1}{2^{\left[\frac{H}{c} N(1-\lambda)\right]}}$$

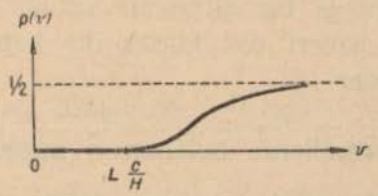
fonksiyonuna gelince, N nin kâfi derecede büyük değerleri için, önceden verilen her sayıdan daha küçüktür. Şu halde, N nin büyük değerleri için, 10^{HN} sayıda takıma dahil kısımların toplam sayısı, $(H/\bar{c})N$ sayıda elementer işaretli kısımların toplam sayısının *ihmâl edilebilen bir kısmını*

teşkil eder; bu keyfiyet, kendilerine farklı takımlar tekabül edecek tarzda, $(H/c)N$ sayıda elementer işaretli 10^{HN} sayıda kısım seçilebileceğini izah eder *. Evvelce de söylediğimiz gibi, böyle bir şifreleme usulü, N nin büyük değerleri için, haberi çok küçük bir hata ihtimali ile almayı temin eder.

Evvelki formüllerde bulunan $\bar{c}$ sayısı c den daha büyük ise, 10^{HN} sayıda takımlara dahil $(H/\bar{c})N$ sayıda harfli kısımların toplam sayısı, mümkün kısımların toplam sayısı olan $M = 2^{\frac{H}{\bar{c}}N}$ den daha büyük olacaktır. Bu halde, alınan haberin çözülmesinde yapılan hata ihtimalinin istenildiği kadar küçük olacak tarzda, birim zamanda $v = L(\bar{c}/H)$ hızı ile bir ulaştırmayı mümkün kılan, $(H/\bar{c})N$ sayıda elementer işaretli 10^{HN} sayıda şifreli yazı bulmak mümkündür. Başka ifade ile eğer ulaştırma hızı zaman biriminde $L \cdot (\bar{c}/H)$ harfden fazla ise, ulaşttırılan her harfin çözülmesindeki hata ihtimali, kullanılan şifre ne olursa olsun, (ulaştırma hızına tâbi olan ve bununla beraber artan) bir sonlu p sayısından daha büyük olacaktır. Basitleştirmek maksadı ile, ulaştırılan haberin bir ikili alfabede şifrelenmiş olduğunu farz edelim; «en iyi» şifreleme usulü vasıtası ile ulaştırılan bir harfin çözülmesinde yapılacak minimum hata ihtimali, v ulaştırma

*) Kendilerine tekabül eden takımlar farklı olacak tarzda bu şifreli yazıların seçilmesi meselesini cevaplandırmamış bulunuyoruz, esasen bu seçimi yapabilmek için (tekmil ulaştırma hatlarında muteber olan) henüz genel metod mevcut değildir.

hızının fonksiyonu olarak şekil 9 da gösterilmiştir. Burada SHANNON teoremi, $v < L \cdot (c/H)$ için $p = p(v)$ eğri-



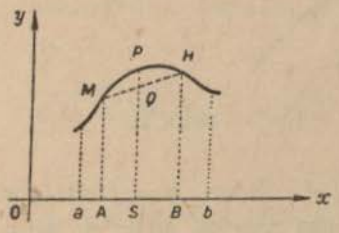
Şekil: 9

sinin apsisler ekseni ile üstüste gelmesine tekabül eder (başka ifade ile, böyle bir ulaştırma hızı ile istenildiği kadar küçük bir hata ihtimalinin elde edilmesi mümkündür), halbuki $v > L \cdot (c/H)$ için bu ihtimal üniform tarzda artar; $v \to \infty$ olduğu vakit hata ihtimalinin 1/2 ye yaklaşması keyfiyeti, herhangi bir haberin büyük bir ulaştırma hızı ile ulaştırmanın pratikte mümkün olmaması ile izah olunur; bu halde, ulaştırılan haberi tesadüfi olarak çözmeye mecbur olacak ve bu takdirde iki harfli bir alfabe için hata ihtimali olan p, 1/2 ye eşit olacaktı.

I. İLÂVE

KONVEKS FONKSİYONLARIN ÖZELİKLERİ

$y = f(x)$ fonksiyonunun gösterdiği eğrinin (a, b) aralığındaki tekmil MN yayı, buna tekabül eden MN kirişi üzerinde bulunursa, $f(x)$ fonksiyonu (a, b) aralığında *konveks*'tir denir * (Şek. 10). Böyle fonksiyonlara misâl

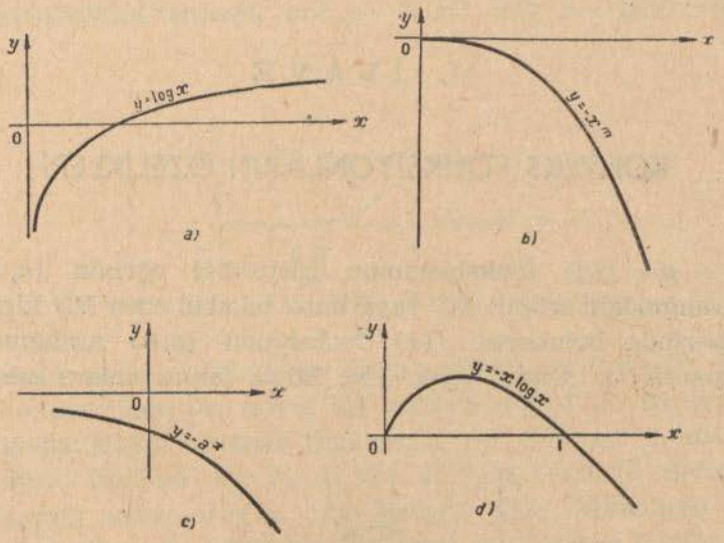


Şekil: 10

olarak, kendi tarif aralığı olan tekmil $(0, +\infty)$ aralığında $y = \log x$ fonksiyonunu, ($m > 1$ olmak üzere) aynı aralıkta $y = -x^m$ fonksiyonunu, $(-\infty, +\infty)$ aralığında $y = -a^x$ fonksiyonunu ve nihayet $(0, \infty)$ aralığında $y = -x \log x$ fonksiyonunu verebiliriz (Şek. 11, *a-d*).

---

*) Yüksek matematikte (hususî olarak bu ilâvede incelenen) geniş bir fonksiyon sınıfının konveksliğini belirlemek için bir kriter mevcuttur. Bir fonksiyonun y'' ikinci türevi negatif ise fonksiyon konvekstir.



Şekil: 11

TEOREM 1. — *Eğer* $y=f(x)$ *fonksiyonu* (a,b) *aralığında konveks ise ve eğer* x_1 ve x_2 *değişkenin bu aralıktaki iki değeri (yani* $a<x_1<x_2<b$*) ise*

$$\frac{f(x_1)+f(x_2)}{2} < f\left(\frac{x_1+x_2}{2}\right) \qquad (1)$$

dir.

ISPAT. — Şekil 10 da, $OA=x_1$ ve $OB=x_2$ olsun; şu halde $AM=f(x_1)$, $BN=f(x_2)$ dir. AB nin orta noktası S olmak üzere $OS=(x_1+x_2)/2$ ve bunun neticesi olarak $SP=f[(x_1+x_2)/2]$ yazılır. Bundan başka,

$ABNM$ yamuğunda kenarların orta noktalarını birleştiren SQ doğru parçası, AM ve BN tabanlarının toplamının yarısına eşit olduğundan $SQ = [f(x_1 + f(x_2)]/2$ dir. Fakat konveks fonksiyonların tarifi mucibince, MN kirişinin Q orta noktası MN yayının üstünde bulunur; buradan

$$\frac{f(x_1) + f(x_2)}{2} < f\frac{x_1 + x_2}{2}$$

yazılır; bu ise teoremi ispat eder.

Misaller.

a) $y = \log x$

$$(\log x_1 + \log x_2)/2 < \log [(x_1 + x_2)/2]$$

yahut

$$\log \sqrt{x_1 x_2} < \log [x_1 + x_2)/2]$$

nihayet

$$\sqrt{x_1 x_2} < (x_1 + x_2)/2$$

bulunur : *iki farklı pozitif sayının geometrik ortalaması, aritmetik ortalamasından küçüktür.*

b) $y = -x^m$, $m > 1$. Burada

$$-(x_1^m + x_2^m)/2 < -(x_1 + x_2)^m/2$$

yahut

$$(x_1^m + x_2^m)/2 > [(x_1 + x_2)/2]^m,$$

$$[(x_1^m + x_2^m)/2]^{1/m} > (x_1 + x_2)/2$$

elde edilir.

$$\left(\frac{a_1^m + a_2^m + \cdots + a_k^m}{k}\right)^{1/m}$$

ifadesi $a_1, a_2 \cdots, a_k$ sayılarının m yinci kuvvetlerinin aritmetik ortalamasının m yinci kökünü gösterir. Şu halde elde edilen netice şöyle ifade edilebilir: *iki farklı pozitif sayının m yinci kuvvetlerinin aritmetik ortalamasının m yinci kökü, $m > 1$ halinde bu sayıların aritmetik ortalamasından büyüktür.*

c) $y = -x \log x$. 1. teoremden

$$-(x_1 \log x_1 + x_2 \log x_2)/2 < [-(x_1 + x_2)/2] \log (x_1 + x_2)/2$$

yahut

$$(-x_1 \log x_1)/2 - (-x_2 \log x_2)/2 < [-(x_1 + x_2)/2] \log [(x_1 + x_2)/2]$$

bulunur; bu neticeyi II. Bölümde iki defa kullandık.

1. teoremdeki eşitsizlik aağşıdaki tarzda genelleştirilebilir:

TEOREM 2. — *Eğer $y = f(x)$, (a, b) aralığında bir konveks fonksiyon ise, eğer x_1 ve x_2 bu aralığın herhangi iki sayısı ise, ve eğer p ve q toplamı 1 olan herhangi iki pozitif sayı ise*

$$pf(x_1) + qf(x_2) < f(x_1 + qx_2) \qquad (2)$$

dir. $p = q = 1/2$ için 2. teorem 1. teoremi verir.

ISPAT. — Evvelâ şuna işaret edelim: *M ve N koordinatları (x_1, y_1) ve (x_2, y_2) olan iki nokta ve Q, MN doğru*

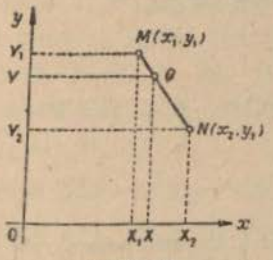
parçasını $MQ/QN = q/p$ $(p+q=1)$ *oranında bölen bir noktası ise* Q *noktasının koordinatları* $px_1 + qx_2$ *ve* $py_1 + qy_2$ *dir.* Gerçekten, X_1, X_2, X ve Y_1, Y_2, Y ile M, N, Q noktalarının koordinat eksenleri üzerindeki izdüşümlerini gösterelim (Şek. 12). X ve Y noktaları X_1X_2 ve Y_1Y_2 doğru parçalarını p/q oranında böler. Buradan şu netice çıkarılır * :

$$OX = OX_1 + X_1X = x_1 + q(x_1 - x_2)$$
$$= (1-q)x_1 + qx_2 = px_1 + px_2$$

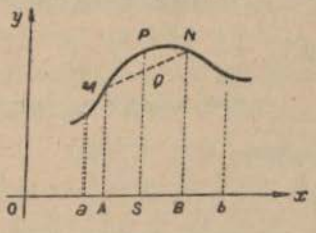
ve

$$OY = OY_2 + Y_2Y = y_2 + p(y_1 - y_2)$$
$$= (1-p)y_2 + py_1 = py_1 + qy_2 .$$

Şimdi yeniden $y = f(x)$ konveks fonksiyonunun gösterdiği eğriyi inceliyelim (Şek. 13); $OA = x_1$, $OB = x_2$, $AM = f(x_1)$, $BN = f(x_2)$ olsun MN doğru parçasını $MN/QN = p/q$ oranında bölen Q noktasının koordinatlarının $px_1 + qx_2$ ve $py_1 + qy_2$ olduğunu gördük; şu



Şekil: 12



Şekil: 13

*) *Şekil* 12 de x_1, x_2 ve y_1, y_2 koordinatları pozitiftir (bu, bizi ilgilendiren haldir). Diğer hallerin incelenmesini okuyucuya bırakıyoruz.

halde $SQ = pf(x_1) + qf(x_2)$ ve $SP = f(px_1 + qx_2)$ dir (gerçekten $OS = px_1 + qx_2$ dir). Fakat $y = f(x)$ fonksiyonu konveks olduğundan Q noktası P noktasının altında bulunur; şu halde

$$pf(x_1) + qf(x_2) < f(px_1 + qx_2)$$

dir; bu ise istenileni ispat eder *.

Misaller :

a) $y = \log x$. Bu halde (2) eşitsizliği

$$p \log x_1 + q \log x_2 < \log (px_1 + xq_2)$$

verir; buradan

$$x_1^p \cdot x_2^q < px_1 + xq_2, \quad p + q = 1$$

bulunur.

b) $y = -x^m$.

$$-px_1^m - qx_2^m < -(px_1 + qx_2)^m$$

yahut

$$px_1^m + qx_2^m > (px_1 + qx_2)^m, \quad p + q = 1$$

elde edilir.

---

* MN doğru parçasının *her* noktasının koordinatlarının $(px_1 + qx_2, py_1 + qy_2)$ şeklinde yazılabileceğini görmek kolaydır; burada $p > 0$, $q > 0$, $p + q = 1$ dir. Şu halde (2) eşitsizliği MN kirişinin $y = f(x)$ eğrisinin altında bulunduğunu ifade eder, yani bu eşitsizlik fonksiyonunun konvekslik tarifine denektir.

c) $Y = -x \log x$. Bu halde

$$-px_1 \log x_1 - qx_2 \log x_2 <$$
$$< -(px_1 + qx_2) \log (px_1 + qx_2), \quad p + q = 1$$

bulunur.

1. teorem şu tarzda da genelleştirilebilir:

TEOREM 3. — *Eğer* $y = f(x)$, (a, b) *aralığında konveks bir fonksiyon ise ve eğer* $x_1, x_2, \cdots, x_k$ *bu aralığın hepsi eşit olmayan* k *tane değeri ise*

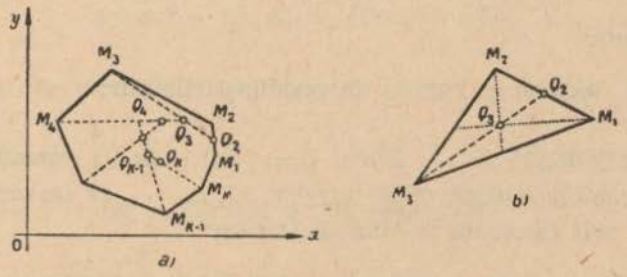
$$\frac{f(x_1) + f(x_2) + \cdots + f(x_k)}{k} <$$
$$< f\left(\frac{x_1 + x_2 + \cdots + x_k}{k}\right) \qquad (3)$$

dır (bu, Jensen eşitsizliğinin özel halidir).

$k = 2$ için teorem 3 tekrar teorem 1 i verir.

ISPAT. — Evvelâ, geometride ve analitik geometride çoğu zaman kullanılan bir anlamın tarifini yapalım. k köşeli $M_1 M_2 M_3 \cdots M_k$ gibi herhangi bir çokköşeli gözönüne alalım (Şek. 14, a); Q_2, M_1M_2 kenarının orta noktası $(M_1Q_2 / Q_2M_2 = \frac{1}{2}/\frac{1}{2})$; Q_3, M_3Q_2 doğru parçasını 2/1 oranında bölen nokta $(M_3Q_3/Q_3Q_2 = \frac{2}{3}/\frac{1}{3})$; Q_4, M_4Q_3 doğru parçasını 3/1 oranında bölen nokta $(M_4Q_4/Q_4Q_3 = \frac{3}{4}/\frac{1}{4})$ $\cdots$; ve nihayet Q_k, M_kQ_{k-1} doğru parçasını $(k-1)/1$ oranında bölen nokta $(M_kQ_k/Q_kQ_{k-1} = \frac{(k-1)}{k}/\frac{1}{k})$ olsun Q_k

noktasına $M_1, M_2, \ldots, M_k$ noktalarının *ağırlık merkezi* denir. Bir $M_1M_2M_3$ üçgeni halinde (Şek. 14, *b*), Q_3 *noktası kenar ortayların kesişme noktası* ile üstüste gelir; gerçekten, bu



Şekil: 14

takdirde, Q_2 noktası M_1M_2 kenarının ortası, M_3Q_3 bir kenar ortay ve Q_3, bu doğru parçasını $M_3Q_3/Q_3Q_2 = 2/1$ oranında bölen noktadır: bu, üçgenin üç kenar ortayının kesişme noktasıdır.

Eğer $M_1, M_2, \cdots, M_k$ *köşelerinin koordinatları* (x_1, y_2), $(x_2, y_2), \cdots, (x_k, y_k)$ *ise,* Q_k *ağırlık merkezinin koordinatlarının*

$$\frac{x_1 + x_2 + \cdots + x_k}{k} \text{ ve } \frac{y_1 + y_2 + \cdots + y_k}{k}$$

olduğunu gösterelim. Gerçekten, teorem 2 nin ispatının başlangıcında faydalanılan özelik mucibince $Q_2, Q_3, Q_4, \cdots$, ve nihayet Q_k noktalarının koordinatları

$$Q_2 - \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right),$$

$$Q_3 - \left(\frac{2}{3}\cdot\frac{x_1+x_2}{2}+\frac{1}{3}x_3,\ \frac{2}{3}\cdot\frac{y_1+y_2}{2}+\frac{1}{3}\cdot y_3\right)$$

yahut $\left(\dfrac{x_1+x_2+x_3}{3},\ \dfrac{y_1+y_2+y_3}{3}\right)$,

$$Q_4 - \left(\frac{3}{4}\cdot\frac{x_1+x_2+x_3}{3}+\frac{1}{4}\cdot x_4,\ \frac{3}{4}\cdot\frac{y_1+y_2+y_3}{3}+\frac{1}{4}\cdot y_4\right)$$

yahut $\left(\dfrac{x_1+x_2+x_3+x_4}{4},\ \dfrac{y_1+y_2+y_3+y_4}{4}\right)$

. . . . . .

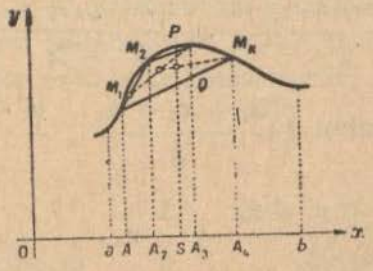
$$Q_k - \left(\frac{(k-1)}{k}\cdot\frac{x_1+x_2+\cdots+x_{k-1}}{k-1}+\frac{1}{k}\cdot x_k,\ \frac{(k-1)}{k}\cdot\frac{y_1+y_2+\cdots+y_{k-1}}{k-1}+\frac{1}{k}\cdot y_k\right)$$

yahut

$$\left(\frac{x_1+x_2+\cdots+x_{k-1}+x_k}{k},\ \frac{y_1+y_2+\cdots+y_{k-1}+y_k}{k}\right)$$

dır.

$y=f(x)$ konveks fonksiyonunu yeniden gözönüne alalım. Bu fonksiyonun gösterdiği egerinin $M_1, M_2, \cdots, M_k$ gibi k tane noktasını düşünelim (Şek. 15). Fonksiyon konveks olduğundan, $M_1M_2\cdots M_k$ çokgeni konveks olacak ve $y=f(x)$ eğr'sinin tamamen altında bulunacaktır. Eğer



Şekil: 15

$M_1, M_2, \cdots, M_k$ noktalarının apsisleri $x_1, x_2, \cdots, x_k$ ise bu noktaların ordinatları $f(x_1), f(x_2), \cdots, f(x_k)$ olup bu noktaların ağırlık merkezinin koordinatları

$$\frac{x_1 + x_2 + \cdots + x_k}{k} \text{ ve } \frac{f(x_1) + f(x_2) + \cdots + f(x_k)}{k}$$

dır; buradan

$$SQ = \frac{f(x_1) + f(x_2) + \cdots + f(x_k)}{k},$$

$$OS = \frac{x_1 + x_2 + \cdots + x_k}{k}$$

ve

$$SP = f\left(\frac{x_1 + x_2 + \cdots + x_k}{k}\right).$$

yazılır (bak. Şek. 15). Fakat, k köşeli konveks bir çokgenin ağırlık merkezi bu çokgenin içinde bulunur (bu, ağırlık merkezinin tarifi neticesidir); şu halde Q noktası P noktasının altında bulunur ve

$$\frac{f(x_1)+f(x_2)+\cdots+f(x_k)}{k} < f\left(\frac{x_1+x_2+\cdots+x_k}{k}\right),$$

olur, bu ise gösterilmesi istenilen neticedir.

Elde ettiğimiz neticeler $M_1, M_2, \cdots, M_k$ noktalarından (hepsi olmamak şartiyle) bazıları üstüste bulunduğu takdirde de muteberdir (başka ifade ile, $x_1, x_2, \cdots, x_k$ sayılarından bazılarının eşit olması halidir); bu takdirde k kenarlı çokgen daha az sayıda kenarlı bir çokgen olur.

Misaller:

a) $y = \log x$. Teorem 3 den

$$\frac{\log x_1 + \log x_2 + \cdots + \log x_k}{k} < \log \frac{x_1+x_2+\cdots+x_k}{k},$$

yahut

$$\sqrt[k]{x_1 x_2 \cdots x_k} < \frac{x_1+x_2+\cdots+x_k}{k}$$

bulunur.

Hepsi birbirine eşit olmayan k sayıda pozitif sayıların geometrik ortalaması bu sayıların aritmetik ortalamasından küçüktür (geometrik ortalama ve aritmetik ortalama teoremi).

b) $y = x^m$, $m > 1$. Bu halde

$$\frac{x_1^m + x_2^m + \cdots + x_k^m}{k} < \left(\frac{x_1+x_2+\cdots+x_k}{k}\right)^m$$

yahut

$$\left(\frac{x_1^m + x_2^m + \cdots + x_k^m}{k}\right)^{1/m} > \frac{x_1 + x_2 + \cdots + x_k}{k}$$

elde edilir.

Hepsi birbirine eşit olmayan k sayıda pozitif sayıların m yinci kuvvetlerinin aritmetik ortalamasının m yinci kökü bu sayıların aritmetik ortalamasından küçüktür $(m > 1)$.

c) $y = -x \log x$. Bu halde teorem 3,

$$-\frac{x_1 \log x_1 + x_2 \log x_2 + \cdots + x_k \log x_k}{k} <$$

$$< -\frac{x_1 + x_2 + \cdots + x_k}{k} \log \frac{x_1 + x_2 + \cdots + x_k}{k} \quad (4)$$

verir.

Nihayet, torem 2 ve teorem 3 ü geneşşeltiren aşağıdaki teoremi gösterelim:

TEOREM 4. — $y = f(x)$ *bir* (a, b) *aralığında bir konveks fonksiyon ve,* $x_1, x_2, \cdots, x_k$ *bu aralığın k tane sayısı olsun; eğer* $p_1, p_2, \cdots, p_k$, *toplamı* 1 *e eşit olan k tane pozitif sayı ise*

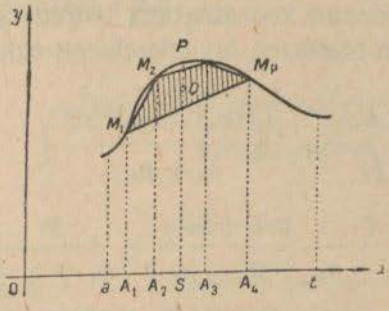
$$p_1 f(x_1) + p_2 f(x_2) + \cdots + p_k f(x_k) <$$

$$< f(p_1 x_1 + p_2 x_2 + \cdots + p_k x_k) \quad (5)$$

dir (genel halde, Jansen eşitsizliği).

Bu teorem $k = 2$ için teorem 2 yi, ve $p_1 = p_2 = \cdots = p_k = 1/k$ için teorem 3 ü verir.

İSPAT. — Yine $y = f(x)$ konveks fonksiyonun gösterdiği eğriyi gözönüne alalım ve bu eğri içine köşelerinin koordinatları (x_1, y_1), (x_2, y_2), $\cdots$, (x_k, y_k) olan k kenarlı $M_1, M_2, \cdots, M_k$ çokgenini çizelim (Şek. 16).



Şekil: 16

$$M_1Q_2/Q_2M_2 = \frac{p_2}{p_1 + p_2} \Big/ \frac{p_1}{p_1 + p_2}$$

olmak üzere çokgenin M_1M_2 kenarı üzerinde bir Q_2 noktasını;

$$M_3Q_3/Q_3Q_2 = \frac{p_3}{p_1 + p_2 + p_3} \Big/ \frac{p_1 + p_2}{p_1 + p_2 + p_3}$$

olmak üzere M_3M_2 doğru parçasının Q_3 noktasını;

$$M_4Q_4/Q_4Q_3 = \frac{p_4}{p_1 + p_2 + p_3 + p_4} \Big/ \frac{p_1 + p_2 + p_3}{p_1 + p_2 + p_3 + p_4}$$

olmak üzere M_4Q_3 doğru parçasının Q_4 noktasını; ve nihayet

$$M_kQ/QQ_{k-1} = p_k/(p_1 + p_2 + \cdots + p_{k-1})$$

olmak üzere M_kQ_{k-1} doğru parçasının Q noktasını gözönüne alalım (eğer $p_1 = p_2 = \cdots = p_k = 1/k$ ise Q noktası $M_1M_2, \cdots, M_k$ çokgeninin ağırlık merkezidir). $Q_2, Q_3, Q_4 \cdots, Q$ noktalarının koordinatları teorem 2 nin ispatının başında verilen özelikten faydalanılarak elde edilir :

$$Q_2 - \left(\frac{p_1x_1 + p_2x_2}{p_1 + p_2}, \frac{p_1f(x_1) + p_2f(x_2)}{p_1 + p_2}\right)$$

$$Q_3 - \left(\frac{p_1 + p_2}{p_1 + p_2 + p_3} \cdot \frac{p_1x_1 + p_2x_2}{p_1 + p_2} + \frac{p_3}{p_1 + p_2 + p_3} x_3,\right.$$

$$\frac{p_1 + p_2}{p_1 + p_2 + p_3} \cdot \frac{p_1f(x_1) + p_2f(x_2)}{p_1 + p_2} +$$

$$\left. + \frac{p_3}{p_1 + p_2 + p_3} f(x_3)\right)$$

yahut

$$\left(\frac{p_1x_1 + p_2x_2 + p_3x_3}{p_1 + p_2 + p_3} \quad \frac{p_1f(x_1) + p_2f(x_2) + p_3f(x_3)}{p_1 + p_2 + p_3}\right)$$

$$Q_4 - \left(\frac{p_1 + p_2 + p_3}{p_1 + p_2 + p_3 + p_4} \cdot \frac{p_1x_1 + p_2x_2 + p_3x_3}{p_1 + p_2 + p_3} + \right.$$

$$\left. + \frac{p_4}{p_1 + p_2 + p_3 + p_4} x_4,\right.$$

$$\frac{p_1+p_2+p_3}{p_1+p_2+p_3+p_4}\cdot\frac{p_1f(x_1)+p_2f(x_2)+p_3f(x_3)}{p_1+p_2+p_3}+$$

$$+\frac{p_4}{p_1+p_2+p_3+p_4}f(x_4)\Bigg)$$

yahut

$$\Bigg(\frac{p_1x_1+p_2x_2+p_3x_3+p_4x_4}{p_1+p_2+p_3+p_4},$$

$$\frac{p_1f(x_1)+p_2f(x_2)+p_3f(x_3)+p_4f(x_4)}{p_1+p_2+p_3+p_4}\Bigg)$$

. . . . . .

$$Q-\Bigg(\frac{p_1x_1+p_2x_2+\cdots+p_{k-1}x_{k-1}+p_kx_k}{p_1+p_2+\cdots+p_{k-1}+p_k},$$

$$\frac{p_1f(x_1)+p_2f(x_2)+\cdots+p_{k-1}f(x_{k-1})+p_kf(x_k)}{p_1+p_2+\cdots+p_{k-1}+p_k}\Bigg)$$

veyahut

$$[p_1x_1+p_2x_2+\cdots+p_kx_k,$$

$$p_1f(x_1)+p_2f(x_2)+\cdots+p_kf(x_k)]$$

zira $p_1+p_2+\cdots+p_k=1$ dir.

Nihayet *şekil* 16 dan

$$SQ=p_1f(x_1)+p_2f(x_2)+\cdots+p_kf(x_k)$$
$$OS=p_1x_1+p_2x_2+\cdots+p_kx_k,$$
$$SP=f(p_1x_1+p_2x_2+\cdots+p_kx_k)$$

elde edilir.

Fakat Q noktası P noktasının altında bulunduğundan (gerçekten, çokgen $y=f(x)$ eğrisinin tamamen altında bulunur ve Q bu çokgenin bir iç noktasıdır),

$$p_1 f(x_1) + p_2 f(x_2) + \cdots + p_k f(x_k) < \\ < f(p_1 x_1 + p_2 x_2 + \cdots + p_k x_k)$$

bulunur, bu ise gösterilmesi istenilen neticedir *.

Misaller.

a) $y = \log x$. Bu halde

$$p_1 \log x_1 + p_2 \log x_2 + \cdots + p_k \log x_k < \\ < \log (p_1 x_1 + p_2 x_2 + \cdots + p_k x_k)$$

elde ederiz, buradan

$$x_1^{p_1} \cdot x_2^{p_2} \cdots x_k^{p_k} < p_1 x_1 + p_2 x_2 + \cdots + p_k x_k, \\ p_1 + p_2 + \cdots + p_k = 1$$

bulunur (*Geometrik ortalama ve aritmetik ortalama teoreminin genelleştirimesi*).

b) $y = -x^m$, $m > 1$.

---

*) $M_1 M_2 \ldots M_k$ çokgeninin içinde *her* noktanın koordinatlarının $(p_1 x_1 + p_2 x_2 + \ldots + p_k x_k,\ p_1 f(x_1) + p_2 f(x_2) + \ldots + p_k f(x_k))$ şekline konulabileceğini görmek kolaydır; burada $p_1 p_2 \ldots p_k$ toplamı 1 e eşit olan k tane pozitif sayıdır. Şu halde (5) eşitsizliği eğri içine çizilen çokgenin bu eğrinin tamamen altında bulunduğunu ifade eder.

$$-p_1x_1^m - p_2x_2^m - \cdots - p_kx_k^m <$$

$$< -(p_1x_1 + p_2x_2 + \cdots + p_kx_k)^m$$

yahut

$$p_1x_1^m + p_2x_2^m + \cdots + p_kx_k^m >$$

$$> (p_1x_1 + p_2x_2 + \cdots + p_kx_k)^m$$

$$p_1 + p_2 + \cdots + p_k = 1$$

elde edilir.

c) $y = -x \log x$. Teorem 4 den

$$-p_1x_1 \log x_1 - p_2x_2 \log x_2 - \cdots - p_kx_k \log x_k <$$

$$< -(p_1x_1 + p_2x_2 + \cdots + p_kx_k) \cdot$$

$$\cdot \log (p_1x_1 + p_2x_2 + \cdots + p_kx_k), \quad (6)$$

$$p_1 + p_2 + \cdots + p_k = 1$$

bulunur.

Bu ilâve kısmın asıl konusunu, (4) ve (6) eşitsizliklerinin ispatını teşkil eder. İlkinden, k mümkün neticesi olan bir α deneyinin entropisinin, *eşit olarak muhtemel* k neticesi olan α_0 deneyinin $\log k$ entropisinden küçük veya buna eşit olduğu neticesi çıkarılır; üstelik, ancak tekmil neticeler eşit olarak muhtemel ise, yani α deneyi α_0 deneyine identik ise $H(\alpha) = \log k$ olabilir. Gerçekten (4) eşitsizliğinin iki tarafını k ile çarpalım ve

$$x_1 = p(A_1),\ x_2 = p(A_2),\ \cdots,\ x_k = p(A_k)$$

koyalım, burada $A_1, A_2 \cdots, A_k$ ile α deneyinin neticeleri gösterilmiştir $[p(A_1) + p(A_2) + \cdots + p(A_k) = 1$ dir;

$p(A_1), p(A_2), \cdots, p(A_k)$ ihtimallerinin hepsi birbirine eşit değildir]; bu takdirde

$$-p(A_1)\log p(A_1) - p(A_2)\log p(A_2) - \\ - \cdots - p(A_k)\log p(A_k) < \\ < -[p(A_1) + p(A_2) + \cdots + p(A_k)] \cdot \\ \cdot \log \frac{p(A_1) + p(A_2) + \cdots + p(A_k)}{k} = -\log \frac{1}{k} = \log k$$

yahut

$$H(\alpha) < H(\alpha_0)$$

elde ederiz.

(6) eşitsizliği, α deneyinin vukuuna bağlı β deneyinin $H_\alpha(\beta)$ entropisinin, bu deneyin bağlı olmayan $H(\beta)$ entropisinden küçük veya eşit olduğunu göstermek için kullanılabilir. Gerçekten, (6) eşitsizliğinde $p_1 = p(A_1)$, $p_2 = p(A_2)$, $\cdots$, $p_k = p(A_k)$ ve $x_1 = p_{A_1}(B_1)$, $x_2 = p_{A_2}(B_1)$, $\cdots$, $x_k = p_{A_k}(B_1)$ koyalım [$A_1, A_2, \cdots, A_k$ ve $B_1, B_2, \cdots, B_k$ ile α ve β deneylerinin neticeleri gösterilmiş olup $p(A_1) + p(A_2) + \cdots + p(A_k) = 1$ dir);

$$-p(A_1)\,p_{A_1}(B_1)\log p_{A_1}(B_1) - p(A_2)\,p_{A_2}(B_1)\log p_{A_2}(B_1) - \cdots \\ \cdots - p(A_k)\,p_{A_k}(B_1)\log p_{A_k}(B_1) < \\ < -[p(A_1)\,p_{A_1}(B_1) + p(A_2)\,p_{A_2}(B_1) + \cdots + p(A_k)\,p_{A_k}(B_1)] \cdot \\ \cdot \log [p(A_1)\,p_{A_1}(B_1) + p(A_2)\,p_{A_2}(B_1) + \cdots + p(A_k)\,p_{A_k}(B_1)].$$

elde ederiz. Fakat toplam ihtimal formülü mucibince

$$p(A_1)\,p_{A_1}(B_1) + p(A_2)\,p_{A_2}(B_1) + \cdots + p(A_k)\,p_{A_k}(B_1) = p(B_1);$$

dir; buradan son eşitsizlik şu şekilde yazılabilir:

$$-p(A_1)\,p_{A_1}(B_1)\log p_{A_1}(B_1) - p(A_2)\,p_{A_2}(B_1)\log p_{A_2}(B_1) - \cdots$$
$$\cdots - p(A_k)\,p_{A_k}(B_1)\log p_{A_k}(B_1) < -p(B_1)\log p(B_1).$$

$p_{A_1}(B_1) = p_{A_2}(B_1) = \cdots = p_{A_k}(B_1) = p(B_1)$ ise (burada son eşitlik toplam ihtimal formülü neticesidir) eşitsizliğimiz eşitlik şeklini alır. Keza

$$-p(A_1)\,p_{A}(B_2)\log p_{A_1}(B_2) - p(A_2)\,p_{A_2}(B_2)\log p_{A_2}(B_2) - \cdots$$
$$\cdots - p(A_k)\,p_{A_k}(B_2)\log p_{A_k}(B_2) < -p(B_2)\log p(B_2),$$

$$-p(A_1)\,p_{A_1}(B_3)\log p_{A_1}(B_3) - p(A_2)\,p_{A_2}(B_3)\log p_{A}(B_3) - \cdots$$
$$\cdots - p(A_k)\,p_{A_k}(B_3)\log p_{A_k}(B_3) < -p(B_3)\log p(B_3).$$

. . . . . .

$$-p(A_1)\,p_{A_1}(B_j)\log p_{A_1}(B_j) - p(A_2)\,p_{A_2}(B_j)\log p_{A_2}(B_j) - \cdots$$
$$\cdots - p(A_k)\,p_{A_k}(B_j)\log p_{A_k}(B_j) < -p(B_j)\log p(B_j).$$

elde edilir. Tekmil bu eşitsizlikleri toplayarak

$$p(A_1)\,H_{A_1}(\beta) + p(A_2)\,H_{A_2}(\beta) + \cdots + p(A_k)\,H_{A_k}(\beta) < H(\beta)$$

yahut

$$H_\alpha(\beta) < H(\beta)$$

elde edilir. Bu eşitsizlik, α ve β deneyleri müstakil değilseler, yani $p_{A_m}(B_n) \neq p(B_n)$ olacak şekilde m ve n $(1 \leq m \leq k,\ 1 \leq n \leq j)$ sayıları mevcut ise, vakidir. Eğer α ve β deneyleri müstakil iseler $H_\alpha(\beta) = H(\beta)$ dir.

Bölüm II, paragraf 3 ün sonunda biraz daha genel olan

$$H_{\alpha\gamma}(\beta \leq H_{\gamma}(\beta)$$

eşitsizliğini kullandık (γ deneyinin 1 ihtimali ile vuku bulan ancak bir neticesi olduğu farzedilirse, bu eşitsizlik $H_{\alpha}(\beta) \leq H(\beta)$ eşitsizliğini verir). Bu eşitsizliği $H_{\alpha}(\beta) \leq H(\beta)$ eşitsizliğinden hareket ederek göstermek kolaydır. Gerçekten, γ deneyinin neticelerini $C_1, C_2, \cdots, C_k$ ile gösterelim; $\alpha^{(1)}$ ve $\beta^{(1)}$ ise $A_1^{(1)}, A_2^{(1)}, \cdots, A_k^{(1)}$ ve $B_1^{(1)}, B_2^{(1)}, \cdots, B_i^{(1)}$ neticelerinin ihtimalleri sıra ile $p(A_1^{(1)}) = p_{C_1}(A_1)$, $p(A_2^{(1)}) = p_{C_1}(A_2)$, $\cdots$, $p(A_k^{(1)}) = p_{C_1}(A_k)$ ve $p(B_1^{(1)}) = p_{C_1}(B_1)$, $p(B_2^{(1)}) = p_{C_1}(B_2)$, $\cdots$, $p(B_i^{(1)}) = p_{C_1}(B_i)$ olsun. Biraz evvel gösterilen neticeye göre

$$H_{\alpha^{(1)}}(\beta^{(1)}) \leq H(\beta^{(1)})$$

dir. Fakat

$$\begin{aligned} H(\beta^{(1)}) = &- p(B_1^{(1)}) \log p(B_1^{(1)}) - p(B_1^{(1)}) \log p(B_2^{(1)}) - \cdots \\ &\cdots - p(B_j^{(1)}) \log p(B_j^{(1)}) = - p_{C_1}(B_1) \log p_{C_1}(B_1) \\ &- p_{C_1}(B_2) \log p_{C_1}(B_2) - \cdots - p_{C_1}(B_i) \log p_{C_1}(B_i) \\ &= H_{C_1}(\beta) \end{aligned}$$

ve

$$\begin{aligned} H_{\alpha^{(1)}}(\beta^{(1)}) = p(A_1^{(1)}) H_{A_1^{(1)}}(\beta^{(1)}) + p(A_2^{(1)}) H_{A_2^{(1)}}(\beta^{(1)}) + \cdots \\ \cdots + p(A_k^{(1)}) H_{A_k^{(1)}}(\beta^{(1)}) \end{aligned}$$

buradan

$$\begin{aligned} H_{A_1^{(1)}}(\beta^{(1)}) = &- p_{A_1^{(1)}}(B_1^{(1)}) \log p_{A_1^{(1)}}(B_1^{(1)}) - \\ &- p_{A_1^{(1)}}(B_2^{(1)}) \log p_{A_1^{(1)}}(B_2^{(1)}) - \cdots - p_{A_1^{(1)}}(B_j^{(1)}) \log p_{A_1^{(1)}}(B_j^{(1)}). \end{aligned}$$

$$\begin{aligned} H_{A_2^{(1)}}(\beta^{(1)}) = &- p_{A_2^{(1)}}(B_1^{(1)}) \log p_{A_2^{(1)}}(B_1^{(1)}) - \\ &- p_{A_2^{(1)}}(B_2^{(1)}) p_{A_2^{(1)}}(B_2^{(1)}) \log - \cdots - p_{A_2^{(1)}}(B_i^{(1)}) \log p_{A_2^{(1)}}(B_j^{(1)}). \end{aligned}$$

. . . . . .

$$H_{A_k^{(1)}}(\beta^{(1)}) = -p_{A_k^{(1)}}(B_1^{(1)}) \log p_{A_k^{(1)}}(B_1^{(1)}) -$$
$$- p_{A_k^{(1)}}(B_2^{(1)}) \log p_{A_k^{(1)}}(B_2^{(1)}) - \cdots - p_{A_k^{(1)}}(B_i^{(1)}) \log p_{A_k^{(1)}}(B_i^{(1)}).$$

yazılır. Şimdi $p_{A_1^{(1)}}(B_1^{(1)})$, $p_{A_1^{(1)}}(B_2^{(1)})$, vs. bağlı ihtimallerini arayalım. Bileşik ihtimaller kanunu mucibince (Bak. Bölüm I, paragraf 3) $p_{A_1^{(1)}}(B_1^{(1)})$ ihtimali $A_1^{(1)}B_1^{(1)}$ ve $A_1^{(1)}$ neticelerinin ihtimallerinin oranına eşittir. Fakat $p(A_1^{(1)}) = p_{C_1}(A_1)$ dir; $A_1^{(1)}B_1^{(1)}$ olayının ihtimaline gelince bu, $p_{C_1}(A_1B_1)$ bağlı ihtimaline eşittir ($A_1^{(1)}$, C_1 olayının vukuuna bağlı A_1 olayının vukuunu, $B_1^{(1)}$ ise, aynı şart altında B_1 olayının vukuunu gösterir; şu halde $A_1^{(1)}B_1^{(1)}$, aynı şart altında A_1B_1 in vukuunu gösterir). Fakat, bileşik ihtimaller kanunu mucibince $p_{C_1}(A_1B_1) = p_{C_1}(A_1)p_{C_1A_1}(B_1)$ dir; buradan

$$p_{A_1^{(1)}}(B_1^{(1)}) = \frac{p(A_1^{(1)}B_1^{(1)})}{p(A_1^{(1)})} = \frac{p_{C_1}(A_1)p_{C_1A_1}(B_1)}{p_{C_1}(A_1)} = p_{C_1A_1}(B_1).$$

elde edilir. Keza

$$p_{A_1^{(1)}}(B_1^{(1)}) = p_{C_1A_1}(B_2),\ p_{A_1^{(1)}}(B^{(1)}) = p_{C_1A_1}(B_3),\ldots$$
$$\ldots p_{A_1^{(1)}}(B_i^{(1)}) = p_{C_1A_1}(B_i).$$

yazılır. Buradan

$$H_{A_1^{(1)}}(\beta^{(1)}) = -p_{C_1A_1}(B_1)\log p_{C_1A_1}(B_1) - p_{C_1A_1}(B_2)\log p_{C_1A_1}(B_2) - \cdots$$
$$\cdots - p_{C_1A_1}(B_i)\log p_{C_1A_1}(B_i) = H_{C_1A_1}(\beta)$$

ve benzer tarzda

$$H_{A_2^{(1)}}(\beta^{(1)}) = H_{C_1A_2}(\beta),\ H_{A_3^{(1)}}(\beta^{(1)}) = H_{C_1A_3}(\beta),\ldots,\ H_{A_k^{(1)}}(\beta^{(1)}) = H_{C_1A_k}(\beta)$$

bulunur. Şu halde [burada $p(A_1^{(1)}) = p_{C_1}(A_1)$, $p(A_2^{(1)}) = p_{C_1}(A_2)$, $\ldots$, $p(A_k^{(1)}) = p_{C_1}(A_k)$ olduğunu hatırlatarak]

$$H_{\alpha}^{(1)}(\beta^{(1)}) = p_{C_1}(A_1) H_{C_1A_1}(\beta) + p_{C_1}(A_2) H_{C_1A_2}(\beta) + \cdots$$
$$\cdots + p_{C_1}(A_k) H_{C_1A_k}(\beta).$$

elde ederiz.

Şu halde $H^{(1)}(\beta^{(1)}) \leq H(\beta^{(1)})$ eşitliği

$$p_{C_1}(A_1) H_{C_1A_1}(\beta) + p_{C_1}(A_2) H_{C_1A_1}(\beta) + \cdots$$
$$\cdots + p_{C_1}(A_k) H_{C_1A_1}(\beta) \leq H_{C_1}(\beta).$$

yazılabilir. Her iki tarafı $p(C_1)$ ile çarparak ve $p(C_1)p_{C_1}(A_1) = p(C_1A_1)$, $p(C_1)p_{C_1}(A_2) = p(C_1A_2)$, . . ., $p(C_1)p_{C_1}(A_k) = p(C_1A_k)$ eşitlikleri hesaba katılarak

$$p(C_1A_1) H_{C_1A_1}(\beta) + p(C_1A_2) H_{C_1A_2}(\beta) + \cdots + p(C_1A_k) H_{C_1A_k}(\beta) \leq$$
$$\leq p(C_1) H_{C_1}(\beta).$$

yazılır. Aynı tarzda

$$p(C_2A_1) H_{C_2A_1}(\beta) + p(C_2A_2) H_{C_2A_2}(\beta) + \cdots + p(C_2A_k) H_{C_2A_k}(\beta) \leq$$
$$\leq p(C_2) H_{C_2}(\beta).$$

$$p(C_3A_1) H_{C_3A_1}(\beta) + p(C_3A_2) H_{C_3A_2}(\beta) + \cdots + p(C_3A_k) H_{C_3A_k}(\beta) \leq$$
$$\leq p(C_3) H_{C_3}(\beta).$$

. . . . . .

$$p(C_hA_1) H_{C_hA_1}(\beta) + p(C_hA_2) H_{C_hA_2}(\beta) + \cdots + p(C_hA_k) H_{C_hA_k}(\beta) \leq$$
$$\leq p(C_h) H_{C_h}(\beta).$$

eşitsizlikleri elde edilir. Bütün bu eşitsizlikleri toplayarak, gösterilmesi istenilen

$$H_{\gamma\alpha}(\beta) \leq H_{\gamma}(\beta)$$

elde edilir ($\gamma\alpha$ ve $\alpha\gamma$ deneyleri farklı değillerdir).

II. İLÂVE

BAZI EŞİTSİZLİKLER

Bu ilâvenin konusu $n! = 1 \cdot 2 \cdot 3 \cdots n$ sayısını hesaplamayı mümkün kılan bazı eşitsizlikleri göstermektir. Fakat önce, e sayısı adı verilen önemli bir sayı tarif etmemize imkân veren olan bazı yardımcı eşitsizliklere ihtiyacımız olacaktır.

1. *n pozitif ve tam herhangi bir sayı olmak üzere*

$$2 \leqslant \left(1 + \frac{1}{n}\right)^n < 3. \qquad (1)$$

dır.

İSPAT. — $k \leq n$ (n tespit edilmiştir) olmak üzere herhangi pozitif ve tam bir k sayısı için

$$1 + \frac{k}{n} \leqslant \left(1 + \frac{1}{n}\right)^k < 1 + \frac{k}{n} + \frac{k^2}{n^2};$$

dir; (1) eşitsizliği bu son eşitsizliğin $k = n$ ye tekabül eden bir özel halini teşkil eder. İspatı rekürans yolu ile yapacağız. $k = 1$ için

$$1 + \frac{1}{n} \leqslant \left(1 + \frac{1}{n}\right)^1 < 1 + \frac{1}{n} + \frac{1}{n^2},$$

dir; bu ise aşikâr olarak doğrudur; şimdi eşitsizliğin k nın belli bir değeri için gösterildiğini farz edelim; bu takdirde

$$\left(1+\frac{1}{n}\right)^{k+1}=\left(1+\frac{1}{n}\right)^{k}\cdot\left(1+\frac{1}{n}\right)\geqslant\left(1+\frac{k}{n}\right)\left(1+\frac{1}{n}\right)=$$

$$=1+\frac{k+1}{n}+\frac{k}{n^2}>1+\frac{k+1}{n}$$

$$\left(1+\frac{1}{n}\right)^{k+1}=\left(1+\frac{1}{n}\right)^{k}\cdot\left(1+\frac{1}{n}\right)<$$

$$<\left(1+\frac{k}{n}+\frac{k^2}{n^2}\right)\left(1+\frac{1}{n}\right)=1+\frac{k+1}{n}+\frac{k^2+k}{n^2}+\frac{k^2}{n^3}=$$

$$=1+\frac{k+1}{n}+\frac{k^2+2k+1}{n^2}-\frac{(k+1)n-k^2}{n^3}<$$

$$<1+\frac{k+1}{n}+\frac{(k+1)^2}{n^2},$$

yazılır, zira $k\leq n$ için $(k+1)n-k^2>0$ dır. Şu halde

$$1+\frac{k+1}{n}<\left(1+\frac{1}{n}\right)^{k+1}<1+\frac{k+1}{n}+\frac{(k+1)^2}{n^2};$$

olduğunu gösterdik; özelik $k=1$ için doğru olduğundan $k\leq n$ olmak üzere, k ne olursa olsun doğrudur [üstelik, $k>1$ için $1+k/n<(1+1/n)^k$ dır].

2. *n pozitif ve tam herhangi bir sayı olmak üzere*

$$\left(1+\frac{1}{n+1}\right)^{k+1}>\left(1+\frac{1}{n}\right)^{n} \qquad (2)$$

ve $n\geqq 2$ ise

$$\left(1+\frac{1}{n}\right)^{n+1} < \left(1+\frac{1}{n-1}\right)^{n}. \qquad (2')$$

dir.

İSPAT. — NEWTON'un binom formülü mucibince,

$$\left(1+\frac{1}{n}\right)^{n} = 1 + C_n^1 \cdot \frac{1}{n} + C_n^2 \cdot \frac{1}{n^2} + \cdots + C_n^{n-1} \cdot \frac{1}{n^{n-1}} + C_n^n \cdot \frac{1}{n^n}$$

$$= 1 + n \cdot \frac{1}{n} + \frac{n(n-1)}{2!} \cdot \frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!} \cdot \frac{1}{n^3} + \cdots$$

$$\cdots + \frac{n(n-1)\cdots[n-(n-2)]}{(n-1)!} \cdot \frac{1}{n^{n-1}} + \frac{n(n-1)\cdots[n-(n-1)]}{n!} \cdot \frac{1}{n^n}$$

$$= 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \cdots$$

$$\cdots + \frac{1}{n!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{n-1}{n}\right)$$

yazılır; keza,

$$\left(1+\frac{1}{n+1}\right)^{n+1} = 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n+1}\right) +$$

$$+ \frac{1}{3!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right) + \cdots$$

$$\cdots + \frac{1}{n!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right)\cdots\left(1-\frac{n-1}{n+1}\right) +$$

$$+ \frac{1}{(n+1)!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right)\cdots\left(1-\frac{n}{n+1}\right).$$

dır. (2) eşitsizliği, bu son iki ifadenin karşılaştırılmasiyle, derhal elde edilir.

Şimdi

$$\left(1+\frac{1}{n}\right)^{n+1} \Big/ \left(1+\frac{1}{n-1}\right)^{n} = \left(\frac{n+1}{n}\right)^{n+1} \Big/ \left(\frac{n}{n-1}\right)^{n} =$$

$$= \frac{(n+1)^{n+1}(n-1)^{n}}{n^{2n+1}} = \left(\frac{n^2-1}{n^2}\right)^{n} \cdot \frac{n+1}{n} =$$

$$= \left(1-\frac{1}{n^2}\right)^{n} \cdot \left(1+\frac{1}{n}\right).$$

oranını teşkil edelim. Fakat $n \geqq 2$ için

$$\left(1-\frac{1}{n^2}\right)^{n} = 1 - n\cdot\frac{1}{n^2} + \frac{n(n-1)}{2!}\cdot\frac{1}{n^4} - \frac{n(n-1)(n-2)}{3!}\cdot\frac{1}{n^6} +$$

$$+ \frac{n(n-1)(n-2)(n-3)}{4!}\cdot\frac{1}{n^8} - \cdots =$$

$$= 1 - \frac{1}{n} + \frac{1}{2!}\,\frac{n-1}{n^3} - \left[\frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdot\frac{1}{n^3} -\right.$$

$$\left. - \frac{1}{4!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\left(1-\frac{3}{n}\right)\cdot\frac{1}{n^4}\right] - \cdots \leqslant$$

$$\leqslant 1 - \frac{1}{n} + \frac{1}{2}\cdot\frac{n-1}{n^3}.$$

ve bunun neticesi olarak

$$\left(1-\frac{1}{n^2}\right)^{n}\left(1+\frac{1}{n}\right) \leqslant \left(1-\frac{1}{n}+\frac{1}{2}\cdot\frac{1}{n^2}-\frac{1}{2}\cdot\frac{1}{n^3}\right)\left(1+\frac{1}{n}\right) =$$

$$= 1 - \frac{1}{2}\cdot\frac{1}{n^2} - \frac{1}{2}\cdot\frac{1}{n^4} < 1.$$

bulunur. Şu halde

$$\left(1+\frac{1}{n}\right)^{n+1} \Big/ \left(1+\frac{1}{n-1}\right)^{n} = \left(1-\frac{1}{n^2}\right)^{n}\left(1+\frac{1}{n}\right) < 1$$

yazılır ve buradan (2′) eşitsizliği elde edilir.

(2) eşitsizliğinden,

$$\left(1+\frac{1}{1}\right)^{1},\ \left(1+\frac{1}{2}\right)^{2},\ \left(1+\frac{1}{3}\right)^{3},\ \ldots,\ \left(1+\frac{1}{n}\right)^{n},\ \ldots$$

dizisinin her sayısının bir sonrakinden daha büyük olduğu görülür. Diğer taraftan, tekmil bu sayılar 3 den küçük olduğundan [bak. (1) eşitsizliği], $n \to \infty$ olduğu vakit, sonsuz olarak artamazlar; şu halde bunlar, aşikâr olarak 2 ile 3 arasında bulunan, belli bir limite yaklaşırlar (dizimizin tekmil sayılarından daha büyük olan). Bu limit değeri e ile gösterilmiştir. e sayısının yaklaşık değeri 2,718281828459045 dir.

Keza, (2′) eşitsizliğinden,

$$\left(1+\frac{1}{2}\right)^{3},\ \left(1+\frac{1}{3}\right)^{4},\ \left(1+\frac{1}{4}\right)^{5},\ \ldots,\ \left(1+\frac{1}{n}\right)^{n+1}$$

dizisinin gittikçe azaldığı görülür (bu sayılardan her biri kendisinden evvelki sayıdan daha küçüktür). Tekmil bu sayılar aşikâr olarak 1 den büyük olduğundan, n sonsuz büyüdüğü vakit, belli bir limite yaklaşırlar; bu limit, dizinin tekmil sayılarından küçük, fakat dizide, n nin kâfi derecede büyük değerine tekabül eden sayılarına istenildiği kadar yakındır. Bu limitin, biraz evvel görülen aynı e sayısı olduğunu görmek kolaydır; gerçekten, ikinci

dizinin bir teriminin birinci dizinin aynı sıradaki terimine oranı olan

$$\left(1+\frac{1}{n}\right)^{n+1} \Big/ \left(1+\frac{1}{n}\right)^{n} = 1+\frac{1}{n}$$

ifadesi, n sonsuz büyüdüğü vakit, 1 e yaklaşır; şu halde iki dizi aynı limite mâliktir.

e sayısı matematikte büyük rol oynar; özel olarak $n!$ sayısının hesabında yer aldığı görülür.

3. *$n \geq 6$ olmak üzere bir tam sayı olsun;*

$$(\frac{1}{2}n)^n > n! > (\frac{1}{3}n)^n \qquad (3)$$

dir.

İSPAT. — Rekürans yolu ile ispatımızı yapalım; $n=1$ için

$$n!=1,\ (\frac{1}{3}n)^n = 1/3, \text{ yani } n! > (\frac{1}{3}n)^n$$

dır. Şimdi n nin belli bir değeri için $n! > (\frac{1}{3}n)^n$ olduğunu farzedelim; buradan,

$$(n+1)! = n! \cdot (n+1) > (\tfrac{1}{3}n)^n \cdot (n+1) =$$

$$= \left(\frac{n+1}{3}\right)^{n+1} \cdot \Big/ \frac{(n+1)}{n \cdot 3} = \left(\frac{n+1}{3}\right)^{n+1} \Big/ \frac{\left(1+\frac{1}{n}\right)}{3} > \left(\frac{n+1}{3}\right)^{n+1}$$

olur, zira (1) eşitsizliğine göre $\frac{1}{3}(1+1/n)^n < 1$ dir. Buradan (yalnız $n \geq 6$ için değil) herhangi pozitif tam n için

$$n! > (\frac{1}{3}n)^n$$

bulunur. Sonra, $n = 6$ için $(\frac{1}{2}n)^n = 3^6 = 729$, $n! = 6! = 720$ dir; yani $(\frac{1}{2}n)^n > n!$ dir. Fakat n nin bir değeri için $n! < (\frac{1}{2}n)^n$ ise aynı zamanda

$$(n+1)! = n!\,(n+1) < (\tfrac{1}{2}n)^n \cdot (n+1) =$$

$$= \left(\frac{n+1}{2}\right)^{n+1} \Big/ \frac{(n+1)}{n^n \cdot 2} = \left(\frac{n+1}{2}\right)^{n+1} \Big/ \frac{\left(1+\frac{1}{n}\right)^n}{2} < \left(\frac{n+1}{2}\right)^{n+1}$$

olacaktır, zira (1) eşitsizliğine göre $\frac{1}{2}(1+1/n)^n > 1$ dir. Şu halde özelik $n \geq 6$ olan her tam n sayısı için doğrudur.

(3) eşitsizliği bizi, 2 ve 3 arasında bulunan bir A sayısının bulunabileceğini düşünmeye sevkeder, şöyle ki, $a < A$ için $(n/a)^n$ kemmiyeti $n!$ den büyük ve $a > A$ için $n!$ den küçük olsun. Gerçekten de bu böyledir; üstelik bu sayı yukarıda tarif ettiğimiz e sayısına eşittir.

4. $n > 6$ *herhangi bir tam sayı ise*

$$\left(\frac{n}{e}\right)^n < n! < n \cdot \left(\frac{n}{e}\right)^n, \qquad (4)$$

dir, buradan $e = 2{,}71828...$ *olup* $n \to \infty$ *için* $(1 + 1/n)^n$ *ifadesinin limitidir.*

(4) eşitsizliğini ispat etmeden evvel, e sayısının neden a sayılarının sınırı olduğunu gösterelim. Gerçekten, $a_1 > e$ ise $n > 6$ için

$$\left(\frac{n}{a_1}\right)^n < \left(\frac{n}{e}\right)^n < n!$$

olacaktır. Bundan başka, $a_2 < e$ ise bir N sayısı bulunabilir. Şöyle ki $n > N$ için $a_2 / \sqrt[n]{1/n}$, e den küçük olsun; bu, $n \to \infty$ için $\sqrt[n]{1/n}$ ifadesinin 1 e yaklaşması neticesidir; gerçekten bu ifadenin logaritması $1/n \log 1/n$ olup $n \to \infty$ için 0 a yaklaşır. Şu halde, $n > N$ için

$$\left(\frac{n}{a_2}\right)^n > n \left(\frac{n}{\frac{a_2}{\sqrt[n]{1/n}}}\right)^n > n \left(\frac{n}{e}\right)^n < n!,$$

olur, bu ise gösterilmesi istenilen neticedir.

İSPAT. — Burada da rekürans yolu ile ispattan faydalanacağız. $n = 1$ için

$$n! = 1, \ (n/e)^n = 1/e \quad \text{yani} \quad n! > (n/e)^n$$

dir. Şimdi n nin belli bir değeri için $n! > (n/e)^n$ olduğunu farzedelim; buradan

$$(n+1)! = n!(n+1) > (n/e)^n \cdot (n+1) =$$

$$= \left(\frac{n+1}{e}\right)^{n+1} \Big/ \frac{(n+1)^n}{n \cdot e} = \left(\frac{n+1}{e}\right)^{n+1} \Big/ \frac{\left(1 + \frac{1}{n}\right)^n}{e} > \left(\frac{n+1}{e}\right)^{n+1}$$

yazılır n ne olursa olsun $(1+1/n)^n/e < 1$ dir. Şu halde n herhangi pozitif tam sayı olmak üzere (yalnız $n > 6$ için değil)

$$n! > (n/e)^n$$

olur. Bundan başka, bir logaritma cetveli yardımı ile

$$7! < 7\cdot(7/e)^7 \quad \text{yani} \quad 6! < (7/e)^7$$

olduğu kolayca gerçeklenir [gerçekten $\log 6! = \log 720 \cong 2{,}857$, $\log (7/e)^7 = 7\,(\log 7 - \log e) \cong 7(0{,}847 - 0{,}434) = 2{,}891$ dir]. Şu halde n nin belli bir değeri için $n! < n\cdot(n/e)^n$ olduğunu farz edersek

$$(n+1)! = n!\,(n+1) < n(n/e)^n\cdot(n+1) =$$

$$= (n+1)\left(\frac{n+1}{e}\right)^{n+1} \Big/ \frac{(n+1)^{n+1}}{n^{n+1}\cdot e} =$$

$$= (n+1)\left(\frac{n+1}{e}\right)^{n+1} \Big/ \frac{\left(1+\frac{1}{n}\right)^{n+1}}{e} < (n+1)\left(\frac{n+1}{e}\right)^{n+1}$$

olacaktır, zira $(1+1/n)^{n+1}/e > 1$ dir. Şu halde özelik $n \geq 7$ olmak üzere n nin herhangi tam değeri için doğrudur.

(4) eşitsizliği daha elverişli şekle sokulabilir. Gerçekten, n nin kâfi derecede büyük değeri için $n!$ sayısı yaklaşık olarak $C\cdot\sqrt{n}(n/e)^n$ ye eşittir, burada C sabit bir sayı olup $\sqrt{2\pi}$ ye eşittir:

$$n! \simeq \sqrt{2\pi n}(n/e)^n$$

(daha kesin ifade ile, $n \to \infty$ için

$$\frac{n!}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}$$

oranı 1 e yaklaşır). Bu neticeye STIRLING *formülü* denir (STIRLING: XVIII inci yüzyılda yaşamış İngiliz matematikçisi); bu formülün ispatı herhangi bir yüksek matematik kitabında bulunabilir (bak. meselâ, Analiz dersleri I; C. J. DE LA VALLÉE POUSSIN; tercüme eden: RATİP BERKER, sahife 206).

TÜRK MATEMATİK DERNEĞİ YAYINLARI

1 — Eşitsizlikler (160 Krş.)
2 — Çokrenk Problemleri (180 Krş.)
3 — Sayılar Teorisinden Problemler (500 Krş.)
4 — Tesadüfi Hareketler (250 Krş.)
5 — Geometrik İspatlarda Hatalar (180 Krş.)
6 — Mekaniğin Matematiğe Bazı Tatbikleri (180 Krş.)
7 — Yalnız Pergelle Yapılan Çizimler (250 Krş.)
8 — Diofant Denklemleri (220 Krş.)
9 — Gruplar Teorisine Giriş (375 Krş.)
10 — Eşitsizliklere Giriş (425 Krş.)
11 — İndüksüyon Metodu (220 Krş.)
12 — Sayılar Teorisine Giriş (550 Krş.)
13 — Geometrik Eşitsizlikler (450 Krş.)
14 — Yalnız Cetvelle Yapılan Çizimler (250 Krş.)
15 — İhtimaller Hesabına Giriş (550 Krş.)
16 — Rasyonel ve İrrasyonel Sayılar (600 Krş.)
17 — Çeşitli Geometriler (750 Krş.)
18 — İndirgemeli Diziler (200 Krş.)
19 — Geometri (Cilt I) (500 Krş.)
20 — Geometri (Cilt II) (750 Krş.)
21 — Geometri (Cilt III) (400 Krş.)
22 — Algoritmalar ve Otomatik Hesap Makinaları (400 Krş.)
23 — Konum Teoremleri (200 Krş.)
24 — Matematiksel Sonsuz (100 Krş.)
25 — Sayılar ve Şekiller (850 Krş.)
26 — Eşdeğer ve Esparçalanabilen Şekiller (300 Krş.)
27 — Oyunlar Teorisine Giriş (300 Krş.)
28 — İhtimaliyet ve İnformasyon (750 Krş.)
29 — Matematikte Endüksiyon ve Benzetme (Cilt I) (750 Krş.)
30 — Matematikte Endüksiyon ve Benzetme (Cilt II) (750 Krş.)

Matematik Derneğinin adresi:
Fen Fakültesi, Matematik Enstitüsü
Vezneciler - İstanbul

Fiatı: 750 Krş.